Università degli Studi di Trento
Facoltà di Scienze
Matematiche, Fisiche e Naturali

Corso di Laurea in Matematica

Eberhard Karls
Universität Tübingen

Studiengang Mathematik

Tesi di Laurea — Diplomarbeit

The concept of dimension: a theoretical and an informal approach

Ester Dalvit

Relatore
Prof. Dr. Domenico Luminati

Betreuer
Prof. Dr. Jürgen Hausen

Anno Accademico 2006–2007

Contents

Introduction **vii**

Introduzione **xi**

Einleitung **xv**

I A survey on dimension 1

1 Dimension of manifolds **3**
1.1 Basic concepts 3
1.2 Tangent space 5
1.3 Sard's Theorem 8
1.4 Some results in Algebraic Topology 15
1.5 Invariance of domain 16

2 Dimension of affine varieties **21**
2.1 Basic concepts 22
2.2 Affine tangent spaces 27
2.3 Derivations 29
2.4 Regular and singular points 36
2.5 Transcendence degree 38
2.6 Morphisms 42
2.7 Krull dimension 45
2.8 Zariski dimension 47
2.9 Krull dimension for rings 49

3 Dimension of metric spaces **53**
3.1 Three concepts of dimension 53
3.2 Some basic theory 57
3.3 The large inductive dimension 60
3.4 Sum theorem for Ind 66

3.5 Subspace and decomposition theorems for Ind 69
3.6 Product theorem for Ind 70
3.7 Covering dimension 71
3.8 Coincidence theorems 80
3.9 The dimension of the Euclidean space 85

II Informal learning of mathematics 89

4 An exhibition about the 4D space 91
4.1 Some basic definitions 92
4.2 Contents of the exhibition 94
4.2.1 Introduction 94
4.2.2 Models 94
4.2.3 Symmetry 97
4.2.4 Slices 97
4.2.5 Dice 98
4.2.6 Flip a cube 98
4.2.7 Polytopes 98
4.3 Flip a cube 99
4.3.1 Animations and images 101
4.3.2 Flip a square 103
4.3.3 Flip a cube 106
4.4 Other animations 110
4.4.1 Folding up a (hyper)cube 110
4.4.2 Projections 112
4.4.3 Hypersphere 117
4.5 Future work 120

Acknowledgments 121

Bibliography 123

List of Figures

1.1 Charts and a chart transformation. 4

4.1 The unfoldings of a cube and a hypercube. 95
4.2 Parallel projections of a cube and a hypercube. 95
4.3 Schlegel diagram of a cube and a hypercube. 95
4.4 A model of the two tori of a hypercube (image by *matematita*). 96
4.5 The rotation axes of the cube in a kaleidoscope. 97
4.6 The cube present in the exhibition. 100
4.7 How to flip a square on the surface of a cube. 101
4.8 Flipping a square on a cube. 103
4.9 Folding up the cube to see the square. 104
4.10 Folding up the cube to see the square. 105
4.11 A square seems a trapezium in the Schlegel diagram. 106
4.12 A square seem a line segment in the Schlegel diagram. 106
4.13 Flipping a cube on a hypercube. 107
4.14 Changing the unfolding of the hypercube. 107
4.15 How to flip a cube on an (unfolded) hypercube. 108
4.16 How to flip a cube on a hypercube (Schegel diagram). 109
4.17 Folding a cube and a hypercube. 110
4.18 How to fold a cube in the 3D space. 111
4.19 How to fold a hypercube in the 4D space. 112
4.20 Projections of a hypercube. 113
4.21 Projections of a cube. 114
4.22 Projections of a hypercube. 115
4.23 A (deformed) sphere inside the solid tori of the hypersphere. . 117
4.24 The sphere deformation in the hypersphere. 118

Introduction

In this work we analyze the concept of dimension from different viewpoints and we describe the preparation of an exhibition about the four dimensional Euclidean space.

The concept of dimension plays a central role in many fields of mathematics, especially in geometry, being one of the first tools used to classify objects. In general, it is important to find some properties that will be used to define and characterize the different classes of classes. The dimension is often one of the first properties that are investigasted.

The most natural class of spaces one can refer to is the class of Euclidean spaces. When one gives a definition of dimension for a class of spaces which contains the Euclidean spaces, it is then important to verify that this definition coincides with the usual one on Euclidean spaces.

The definitions of dimension are often given independently in the different fields of mathematics. It is important to verify that they coincide on spaces that can be considered as subjects of different fields. Nevertheless, even in the cases where the coincidence of different definitions is a well known fact, the proofs are often difficult to be found in the literature.

Even if in mathematics one defines and studies spaces with an arbitrary number of dimensions — the dimension can be a natural or real number or even infinite —, most people have only an intuitive notion of dimension, which derives from the experience. It is then natural to think of objects as having length, width, and height and also distinguish the different nature of linear, plane and solid objects. The relevant dimension of the thread is its length, while to measure a sheet one needs length and height, and for a die also a third dimension is to be considered. In some sense, we consider the thread as one dimensional and the sheet as two dimensional, even if both objects are three dimensional. This is a first "intuitive and informal formalization" of the concept of dimension and is already clear to children. We can exploit this notion to give an idea of the properties of four dimensional objects and, reflecting on similarities and differences between linear, plane and solid objects, it is possible to "jump" to the four dimensions.

The concept of n-dimensional space ($n \geq 4$) is not easy to be illustrated, especially because people often try to find representations through drawings, but it can be very hard to find efficacious representations of objects in a higher dimension and also to interpret and understand them.

Nevertheless, imagination and fantasy of many authors have been fascinated by the four dimensional space and the literature has shown interest towards the concept of dimension since the late 19th century. One of the first works about the concept of dimension and perhaps the most famous, which has inspired many other literary works, is *Flatland: A Romance of Many Dimensions*, published by Edwin A. Abbott in 1884. The author uses the novel to criticize the society of his time, but this work is very interesting also from the mathematical point of view. The protagonist is a square, living in a two dimensional world, which gets in touch with a sphere and makes a journey in the three dimensional world. In this manner the author presents the three dimensional objects from the point of view of the square, which is at first sceptical about the existence of a third dimension. This is perhaps the first time that the general public is invited to reflect on the concept of dimension and on the possibility that other universes with a different number of dimensions can exist or at least can be imagined.

In the same period Charles H. Hinton published two "scientific romances", as he called them, *What is the Fourth Dimension?* and *A Plane World*, where he presents some ideas about the concept of dimension. He also invented the words *ana* and *kata* (from the Greek "up" and "down"), to indicate the two opposite directions in the fourth dimension, the equivalents of left and right, forwards and backwards, and up and down. *Vinn* and *vout* are other words with the same meaning of *ana* and *kata*, coined by the science fiction author Rudy Rucker in his 2002 book *Spaceland*, a story in which a man meets a four dimensional being. Henry More, an English philosopher who lived in the 17th century, had already given a name to the fourth spatial dimension, calling it *spissitude*. He thought that the spiritual beings are four dimensional and they can reduce their extension in the first three dimensions to become larger in the fourth dimension.

Many other short stories and novels dealing with the concept of dimensions have been written. Ian Stewart indagates various concepts of dimensions in his 2001 book *Flatterland*. *The Planiverse* is a novel by Alexander K. Dewdney, written in 1984, which describes a two dimensional world, its physics, biology and chemistry. The novel *Sphereland* was written in 1965 by Dionys Burger as a sequel of *Flatland* and indagates the spherical geometry, to arrive to the concept that our universe is the hypersurface of a hypersphere. Robert A. Heinlein's novel *"–And He Built a Crooked House–"* was written in 1941 and talks about an adventure in a hypercubic house.

The science fiction, both in films and books, often uses the concept of "hyperspace" to give an atmosphere of mistery and suggestion. Exploiting the people's curiosity for these topics makes perhaps possible to draw the attention of the general public to an exhibition on four dimensional objects and different ways of representing and visualizing them. Suggestive ideas like hyperspace are powerful tools to make the visitor be interested in the exhibition and many related topics can be presented and explained, for example projections, symmetries, rotations, polyhedra and polytopes, sections, knots and links. This is the aim of an exhibition which is being prepared by the interuniversitary research centre for the communication and informal learning of mathematics *matematita*. A part of the work for this thesis was done within the project of the exhibition.

Many ideas of this work were inspired by the work of Thomas Banchoff, in particular the animations about the hypercube and the hypersphere. Banchoff uses electronic media to support his mathematical researches and as a powerful tool in mathematical educational. He developed many web-based projects[1] to use computer graphics for education and informal learning of mathematics. Also his book *Beyond the Third Dimension* and the web pages[2] with the same title, which contain a virtual exhibition about mathemtical objects, recall and explain the concept of dimension.

The present work is divided into two parts: in the first one, consisting of three chapters, we present various definitions of dimension, giving attention to different fields of mathematics.

In the first chapter we consider the differential topology and in particular we recall some basic theory about manifolds and the Sard's Theorem. With these tools it can be proved that spheres and Euclidean spaces with different dimensions can not be homeomorphic. Using homology theory, these and other results can be obtained more directly. In particular it can be shown that, if $n \neq m$, then the Euclidean spaces with dimension n and m can not be locally homeomorphic (Theorem on Invariance of Domain). Using the latter results, it can be shown that the dimension of topological manifolds is well defined.

In chapter 2 we pass to consider affine algebraic varieties. Here we investigate the dimension of an algebraic set defined as dimension of the affine tangent space and the correlation between this concept and derivatives. Then

[1] Some of Thomas Banchoff's projects are listed and linked in the web page `http://www.math.brown.edu/TFBCON2003/mathematics/`. Other interesting web pages about some four dimensional objects by Thomas Banchoff and Davide Cervone can be found at `http://www.math.union.edu/~dpvc/math/4D/`.

[2] `http://alem3d.obidos.org/en/`, realized by Banchoff and Cervone.

we recall the concepts of regular and singular points and show the connection of this concept with the dimension. Finally the dimension as transcendence degree and the Krull dimension are presented. It is proved that all these definitions of dimension coincide. To conclude the chapter we outline the proof of the connections between this characterization of the dimension and the stalks of holomorphic functions.

In chapter 3 three different definitions of dimension are presented for metric spaces. It is proved that the different definitions coincide on a large class of metric spaces, and that the n-dimensional Euclidean space has dimension n also in the sense of metric space.

The second part of this work contains a chapter, where we describe the work done for the preparation of an exhibition about some four dimensional objects, which has taken place in Genova during the *Festival della Scienza*, from the 25th October to the 6th November 2007. We designed and developed some interactive animations and animated images to illustrate some models of 4D objects and to explain some topics, like projections and rotations. The animations have been organized in some web pages, which can be found at `http://matematita.science.unitn.it/4d/` and, together with the source code, in the CD–ROM included to this thesis.

Introduzione

In questo lavoro verrà analizzato da vari punti di vista il concetto di dimensione e sarà descritta la preparazione di una mostra sullo spazio euclideo quadridimensionale.

Il concetto di dimensione ha un ruolo centrale in molti campi della matematica, a cominciare dalla geometria, dove è uno dei primi strumenti utilizzati per classificare gli oggetti. In generale si rivela utile cercare quelle proprietà che servono per definire e caratterizzare le diverse classi di spazi; la dimensione è spesso una delle prime proprietà che si considerano.

La classe di spazi più semplice e naturale da considerare è formata dagli spazi euclidei. Quando viene definita la dimensione per una classe di spazi di cui fanno parte anche gli spazi euclidei, è importante verificare che essa coincida con la solita dimensione di spazio euclideo, cioè che $\mathbb{R}^n$ abbia dimensione n in ogni senso.

La dimensione viene spesso introdotta in modo indipendente nei diversi campi della matematica e si pone allora il problema di verificare che queste definizioni coincidano sugli spazi che possono essere considerati oggetto di studio di diverse discipline. Spesso, anche nei casi in cui l'equivalenza delle diverse definizioni è nota, le dimostrazioni non si trovano facilmente in letteratura.

Anche se in matematica si definiscono e si studiano spazi di dimensione arbitraria — la dimensione può essere un qualsiasi numero naturale o reale o anche infinita —, la maggior parte delle persone ha soltanto un'idea intuitiva di dimensione, idea legata alla propria esperienza. È perciò naturale considerare oggetti che abbiano una lunghezza, una larghezza e un'altezza e anche distinguere tra oggetti di natura diversa: oggetti lineari, piani o solidi. Per un filo la dimensione rilevante è la lunghezza, per un foglio di carta lo è anche la larghezza, mentre per un dado bisogna considerare anche una terza dimensione, l'altezza. In qualche senso, il filo è 1-dimensionale e il foglio 2-dimensionale, anche se in realtà entrambi hanno tre dimensioni. Questa nozione intuitiva e informale di dimensione ci appartiene fin da bambini e su di essa si può far leva per far intuire quali possano essere le proprietà di un

oggetto a quattro dimensioni. Si può riflettere su somiglianze e differenze tra oggetti a una, due, tre dimensioni e quindi per analogia compiere il "salto" nello spazio a quattro dimensioni.

Il concetto di spazio n-dimensionale ($n \geq 4$) non è semplice da spiegare, in particolare perché le persone tendono a cercare rappresentazioni grafiche, ma non è facile nè trovare immagini efficaci di oggetti multidimensionali nè interpretarle quando ci vengono proposte.

Nonostante questo, lo spazio quadridimensionale ha esercitato ed esercita un grande fascino sull'immaginazione e sulla fantasia e la letteratura si è interessata al concetto di dimensione già a partire dalla fine del diciannovesimo secolo. Una delle prime opere sul concetto di dimensione, e forse la più famosa, che ha ispirato molti altri racconti e romanzi, è *Flatland: A Romance of Many Dimensions* (*Flatlandia: racconto fantastico a più dimensioni*), pubblicato da Edwin A. Abbott nel 1884. L'autore si serve del racconto per criticare la società del suo tempo, ma il suo lavoro è molto interessante anche dal punto di vista matematico. Il protagonista è un quadrato che vive in un mondo a due dimensioni e che incontra una sfera, con la quale compie un viaggio nel mondo tridimensionale. In questo modo l'autore presenta gli oggetti tridimensionali dal punto di vista del quadrato, all'inizio molto scettico riguardo all'esistenza di una terza dimensione. Questa è forse la prima volta che il grande pubblico è invitato a riflettere sul concetto di dimensione e sulla possibilità che universi con un diverso numero di dimensioni possano esistere o perlomeno essere immaginati.

Nello stesso periodo Charles H. Hinton pubblica i due "racconti scientifici" — così chiamati da lui stesso — *What is the Fourth Dimension?* (*Che cos'è la quarta dimensione?*) e *A Plane World* (*Un mondo piano*), in cui presenta alcune idee sul concetto di dimensione. Egli inventa anche le parole *ana* e *kata* (dal greco "su" e "giù"), per indicare le due direzioni opposte nella quarta dimensione, in aggiunta a destra-sinistra, avanti-indietro e su-giù. *Vinn* e *vout* sono altre parole con lo stesso significato di *ana* e *kata*, coniate dallo scrittore di fantascienza Rudy Rucker nel suo libro *Spaceland*, pubblicato nel 2002, che racconta un incontro fantascientifico tra un uomo e un essere a quattro dimensioni. Già un filosofo inglese del diciassettesimo secolo, Henry More, aveva dato un nome alla quarta dimensione spaziale, chiamandola *spissitude*. Secondo il suo pensiero, le quattro dimensioni sono proprie degli spiriti, che possono diminuire la propria estensione nelle tre dimensioni per acquistare maggior estensione nella quarta.

Sono stati scritti molti altri racconti e romanzi che hanno a che fare con il concetto di dimensione. Ian Stewart parla di diversi concetti di dimensione nel libro *Flatterland*, pubblicato nel 2001. *The Planiverse* è un racconto di Alexander K. Dewdney, scritto nel 1984, che descrive un mondo

bidimensionale, la sua fisica, biologia e chimica. Il racconto *Sphereland*, pubblicato nel 1965 da Dionys Burger, è un seguito di *Flatland* ed esamina la geometria sferica, per arrivare all'idea che il nostro universo è l'ipersuperficie di un'ipersfera. Il racconto di Robert A. Heinlein intitolato *"—And He Built a Crooked House—"* (*La casa ipercubica*) è stato scritto nel 1941 e racconta un'avventura fantascientifica in una casa che si richiude nella quarta dimensione, assumendo la forma di un ipercubo.

La fantascienza, sia nei libri che nei film, usa spesso il concetto di "iperspazio" per creare un'atmosfera di mistero e suggestione. Sfruttando la curiosità del pubblico per questi argomenti suggestivi è forse possibile richiamare visitatori ad una mostra che racconti alcuni oggetti quadridimensionali e vari modi di rappresentarli e visualizzarli. Si possono allora presentare e sviluppare alcuni argomenti riguardanti il concetto di dimensione, come ad esempio proiezioni, simmetrie, rotazioni, poliedri e politopi regolari, sezioni e nodi. Questo è lo scopo di una mostra che il Centro interuniversitario per la comunicazione e l'apprendimento informale della matematica *matematita* sta preparando e nel cui ambito è stata realizzata una parte del lavoro di tesi.

Molte delle idee presenti in questo lavoro, in particolare le animazioni su ipercubi e ipersfere, sono state ispirate dal lavoro di Thomas Banchoff, uno fra i primi ad usare la *computer graphics* sia nella ricerca matematica che nell'insegnamento. Egli ha sviluppato molti progetti sul web[3], sfruttando la grafica tridimensionale per l'insegnamento informale della matematica. Anche il suo libro *Beyond the Third Dimension* (*Oltre la terza dimensione*) e le pagine web[4] con lo stesso titolo, che contengono una mostra virtuale di oggetti matematici, richiamano e cercano di spiegare il concetto di dimensione.

Il lavoro presente è diviso in due parti: nella prima, che contiene tre capitoli, vengono presentate varie definizioni di dimensione, considerando diversi campi della matematica.

Nel primo capitolo viene presa in esame la topologia differenziale e in particolare si richiama la teoria di base delle varietà differenziabili e il teorema di Sard. Con questi strumenti si può dimostrare che sfere e spazi euclidei di dimensioni diverse non possono essere omeomorfi. Con la teoria omologica questi e altri risultati vengono ottenuti più direttamente; in particolare si

[3]Alcuni links ai progetti di Banchoff si possono trovare nella pagina web `http://www.math.brown.edu/TFBCON2003/mathematics/`. Altre pagine riguardanti oggetti quadridimensionali, scritti da Thomas Banchoff e Davide Cervone, si possono visionare all'indirizzo `http://www.math.union.edu/~dpvc/math/4D/`.

[4]`http://alem3d.obidos.org/en/` realizzato da Banchoff e Cervone.

prova che, se $n \neq m$, due spazi euclidei di dimensione n e m rispettivamente non possono essere nemmeno localmente omeomorfi (teorema di invarianza del dominio). Questi risultati assicurano che la dimensione delle varietà topologiche è ben definita.

Nel capitolo 2 vengono considerate le varietà algebriche affini. Dapprima viene esaminata la dimensione degli insiemi algebrici, definita come dimensione dello spazio tangente affine, e la correlazione di questo concetto con le derivazioni. Quindi viene richiamata la nozione di punti regolari e singolari, spiegando come essa abbia a che fare con la dimensione. Infine vengono presentate la dimensione come grado di trascendenza e la dimensione di Krull e viene dimostrato che tutte queste definizioni di dimensione coincidono. Per concludere il capitolo, viene illustrato uno schema di dimostrazione della relazione tra questa caratterizzazione di dimensione e i fasci di funzioni olomorfe.

Nel capitolo 3 vengono introdotte tre diverse definizioni di dimensione per gli spazi metrici. Si dimostra che esse coincidono su un'ampia classe di spazi metrici e che lo spazio euclideo $\mathbb{R}^n$ ha dimensione n anche secondo queste definizioni.

La seconda parte di questo lavoro contiene un capitolo che descrive la preparazione di una mostra su alcuni oggetti quadridimensionali, che è stata allestita a Genova in occasione del *Festival della Scienza*, dal 25 ottobre al 6 novembre 2007. Sono state progettate e sviluppate alcune animazioni interattive e immagini animate per illustrare alcuni modelli di oggetti quadridimensionali e per spiegare qualche argomento correlato, come proiezioni e rotazioni. Le animazioni sono state organizzate in pagine web, disponibili all'indirizzo `http://matematita.science.unitn.it/4d/` e nel CD–ROM allegato alla tesi, dove si trova anche il codice sorgente delle animazioni.

Einleitung

In dieser Arbeit wird der Dimensionalbegriff aus verschiedenen Gesichtspunkte analysiert und die Vorbereitung einer Ausstellung über den vierdimensionalen euklidischen Raum wird beschrieben.

Der Begriff der Dimension spielt eine zentrale Rolle in vielen Beriche der Mathematik, insbesondere in der Geometrie und ist ein der ersten Mittel zur Klassifikation von Objekte. Es ist nämlich wichtig, die Eigenschaften zu untersuchen, die man braucht, um die verschiedene Klassen von Räume zu definieren und kennzeichnen; die Dimension ist oft eine der ersten betrachteten Eigenschaften.

Die einfachste und natürlichste Klasse von Räume, die man betrachten kann, besteht aus den euklidischen Räume. Wenn die Dimension für eine Klasse definiert wird, die auch die euklidische Räume enthält, ist es dann wichtig zu prüfen, dass die neue Definition mit der übliche für euklidische Räume übereinstimmt, d.h. $\mathbb{R}^n$ soll in jedem Sinn die Dimension n besitzen.

Die Dimension wird oft unabhängig in jedem Bereich eingeführt und dann entsteht das Problem, ob die Definitionen über die Räume übereinstimmen, die als Objekte verschiedener Bereiche betrachtet werden können. In vielen Fällen hat diese Frage eine positive Antwort, aber der Beweis ist oft nicht einfach zu finden in der Literatur.

Obwohl in der Mathematik Räume beliebiger Dimension betrachtet werden — die Dimension kann eine beliebige natürliche oder reelle Zahl sein, oder auch unendlich —, meistens haben Leute nur einen intuitiven Begriff der Dimension, der aus der praktischen Erfahrung stammt. Es ist nämlich natürlich, die Länge, die Breite und die Höhe der Gegenstände zu betrachten, und auch unter Sachen verschiedener Natur zu unterscheiden: lineare, flache oder räumliche Sachen. Die bedeutende Dimension eines Fadens ist die Länge, für ein Blatt ist auch die Breite wichtig, und für ein Würfel ist auch eine dritte Dimension zu betrachten, die Höhe. Irgendwie betrachten wir den Faden als eindimensional und das Blatt als zweidimensional, obwohl beide Gegenstände drei Dimensionen besitzen. Diesen intuitiven und informellen Begriff der Dimension besitzen wir schon als Kinder, und er kann benutzt

werden, um eine Idee zu geben, welche Eigenschaften ein vierdimensionales Objekt besitzt. Es ist möglich, über Ähnlichkeiten und Unterschiede unter ein-, zwei- und dreidimensionale Gegenstände zu überlegen und durch Analogie in den vierdimensionalen Raum zu "springen".

Es ist nicht einfach, den Begriff des n-dimensionalen Raumes informell zu erklären, insbesondere weil Leute grafische Darstellungen suchen, aber es ist nicht trivial weder wirksame Darstellungen zu finden noch sie interpretieren und verstehen.

Jedenfalls übt der vierdimensionale Raum einen starken Einfluss auf die Fantasie und die Imagination vieler Schriftsteller und die Literatur hat sich schon seit dem Ende des 19. Jahrhunderts für dieses Thema interessiert. Eine der ersten literarischen Werke über den Begriff der Dimension, und vielleicht die berühmteste, die viele Erzählungen und Romane inspiriert hat, ist *Flatland: A Romance of Many Dimensions*, von Edwin A. Abbott in 1884 veröffentlicht. Der Autor benutzt die Geschichte, um die Gesellschaft seiner Zeit zu kritisieren, aber dieses Werk ist auch aus dem mathematischen Gesichtspunkt interessant: der Protagonist ist ein Viereck, das in einer zweidimensionalen Welt lebt und eine Sphäre trifft, mit der es eine Reise in der dreidimensionalen Welt macht. Der Autor beschreibt die dreidimensionale Objekte aus der Sichtspunkt des Quadrats, das am Anfang skeptisch über die Existenz einer dritten Dimension ist. Die breite Öffentlichkeit ist vielleicht zum ersten Mal eingeladen, über den Begriff der Dimension zu überlegen und über die Möglichkeit der Existenz oder mindestens der Vorstellung eines Universums mit einer anderen Anzahl von Dimensionen.

In der selben Zeit veröfflicht Charles H. Hinton zwei sogennante "scientific romances" *What is the Fourth Dimension?* und *A Plane World*, wo er einige Ideen über den Begriff der Dimension vorstellt. Er erfindet die Wörter *ana* und *kata* (aus der alten griechischen Sprache "auf" und "unter"), um die gegenseitige Richtungen in der vierte Dimension zu kennzeichnen, analog zu links–rechts, vorwärts–rückwärts, oben–unten. *Vinn* und *vout* sind gleichbedeutende Wörter mit *ana* und *kata*, die von dem Science-Fiction-Schriftsteller Rudy Rucker geprägt wurden. In seinem Roman *Spaceland*, veröffentlicht in 2002, erzählt Rucker, wie ein Mann ein vierdimensionales Lebewesen trifft. Schon der Philosoph Henry More, der im 17. Jahrhundert lebte, hatte einen Name zu der vierten Dimension gegeben: *spissitude*. Er dachte, die vier Dimensionen seien von den Geisten besitzt, die die erste drei Dimensionen kleiner machen könnten, um in die vierte sich zu erweitern.

Viele weitere Geschichte und Romane wurden geschrieben, die vom Begriff der Dimension handeln. Ian Stewart erzählt von verschiedenen Vorstellungen der Dimension im Buch *Flatterland*, veröffentlicht in 2001.

The Planiverse ist eine Geschichte von Alexander K. Dewdney, geschrieben in 1984, die eine zweidimensionale Welt, seine Physik, Biologie und Chemie beschreibt. , scritto nel 1984, che descrive un mondo bidimensionale, la sua fisica, biologia e chimica. Die Erzählung *Sphereland*, veröffentlicht in 1965 von Dionys Burger, ist eine Fortsetzung von *Flatland* und betrachtet die sphärische Geometrie, um zur Idee zu kommen, dass unsere Welt die Hyperfläche einer vierdimensionalen Sphäre ist. Die Geschichte *"—And He Built a Crooked House—"*, von Robert A. Heinlein in 1941 veröffentlicht, erzählt ein Science-Fiction-Abenteuer in einem Haus, das in der vierte Dimension zusammenfaltet und die Form eines Hyperkubus annimmt.

Die Science-Fiction, sowohl in Bücher als auch in Filme, benutzt oft den Begriff von "Hyperrraum", um eine geheimnisvolle und eindrucksvolle Stimmung zu schaffen. Es ist vielleicht möglich, die Kuriosität des Publikums auszunutzen, um Besucher zu einer Ausstellung anzulocken, die einige vierdimensionale Gegenstände und deren verschiedene Darstellungen einführt. Man kann dann einige Themen vorstellen, die mit dem Begriff der Dimension verbunden sind, wie zum Beispiel Projektionen, Symmetrien, Rotationen, regelmässige Polyeder und Polytope, Schnitte und Knoten. Das ist das Ziel einer Ausstellung, die wird gerade vorbereitet von dem Universitätszentrum für die Kommunikation und das informelles Lehren der Mathematik *matematita*. Im Rahmen dieses Projektes wurde Teil der vorliegenden Arbeit realisiert.

Viele in dieser Arbeit enthalten Ideen, insbesondere die Computeranimationen über Hyperkuben und Hypersphären, wurden von der Arbeit von Thomas Banchoff inspiriert, einer der ersten Mathematikern, der die Computergrafik sowohl in der Forschung als auch im Lehren benutzt. Er hat mehrere Web-Projekte[5] entwickelt, und benutzt die 3D Grafik für das informelle Lehren der Mathematik. Sein Buch *Beyond the Third Dimension* und die Webseiten [6] mit dem selben Titel, die eine virtuelle Ausstellung über mathematische Objekte enthalten, erinnern und erklären den Begriff der Dimension.

Die vorliegende Arbeit ist in zwei Teile geteilt: in dem ersten, der drei Kapitel enthält, werden Definitionen von Dimension vorgestellt, die von verschiedenen Bereiche der Mathematik untersucht werden.

Im ersten Kapitel wird die Differentialgeometrie betrachtet; insbesondere

[5]Einige Links zu den Banchoffs Projekte sind auf der Seite `http://www.math.brown.edu/TFBCON2003/mathematics/` zu finden. Weitere Webseiten über vierdimensionalen Objekte, von Thomas Banchoff und Davide Cervone, sind auf `http://www.math.union.edu/~dpvc/math/4D/` verfügbar.

[6]`http://alem3d.obidos.org/en/` realisiert von Banchoff und Cervone.

werden die Grundlagen der Theorie für differenzierbare Mannigfaltigkeiten und den Satz von Sard erinnert. Mit diesen Mittel kann man zeigen, dass Sphären und euklidische Räume unterschiedlicher Dimension nicht homöomorph sind. Mit der Homologie werden diese und weitere Ergebnisse unmittelbar erhalten. Insbesondere zeigt man, dass zwei euklidische Räume, bzw. der Dimension n und m, wobei $n \neq m$, auch nicht lokal homöomorph sein können. Diese Resultate sichern, dass die Dimension der Mannigfaltigkeiten wohldefiniert ist.

Im Kapitel 2 werden algebraische affine Varietäten betrachtet. Zuerst wird die Dimension algebraischer Mengen analysiert, die definiert ist als Dimension des affinen Tangentialraumes; dieser Begriff ist mit den Derivationen verbunden. Die Definition von reguläre und singuläre Punkte wird dann erinnert und es wird erklärt, was sie mit der Dimension zu tun hat. Dann werden die Dimension als Transzendenzgrad und die Krulldimension vorgestellt und es wird gezeigt, dass alle Definitionen übereinstimmen. Schliesslich wird eine Beweisskizze der Korrelation zwischen diese Charakterisierungen der Dimension und die Garbe holomorpher Funktionen vorgestellt.

Im Kapitel 3 werden drei Definitionen für metrische Räume eingeführt. Es wird gezeigt, dass sie über eine grosse Klasse von metrische Räume übereinstimmen, und dass der euklidische Raum $\mathbb{R}^n$ auch in diesem Sinn die Dimension n besitzt.

Der zweite Teil dieser Arbeit beschreibt die Vorbereitung einer Ausstellung über einige vierdimensionale Objekte, die in Genova eingerichtet wurde, im Rahmen des *Festival della Scienza*, vom 25. Oktober bis zum 6. November 2007. Einige Animation Bilder und interaktive Computeranimationen wurden geplant und entwickelt, um Modelle von vierdimensionalen Objekte zu erklären und einige Themen wie Projektionen und Rotationen zu erläutern. Die Computeranimationen sind in Webseiten organisiert, die auf `http://matematita.science.unitn.it/4d/` und in der zu dieser Arbeit angelegten CD–ROM zu finden sind. In der CD–ROM ist auch den Quellcode verfügbar.

Part I

A survey on dimension

Chapter 1

Dimension of manifolds

In this chapter we consider the dimension of manifolds, from the point of view of the differential topology and the algebraic topology.

Recall that a topological manifold is a Hausdorff space, which is locally homeomorphic to some $\mathbb{R}^n$ and has a countable basis. The integer n is called the dimension of the manifold. However, the homeomorphism with $\mathbb{R}^n$ is a local property, and one has to check that the n is the same for the whole manifold.

It must also be proved that for each point there can not be two homeomorphisms of some neighbourhoods to Euclidean spaces with different dimension.

We will introduce the notion of regular points for a differentaible manifold, and prove the Sard's Theorem 1.3.5. The latter result will be used to show that spheres with different dimensions can not be homeomorphic.

Furthermore, using the homology theory, it can be proved that manifold with different dimensions are not locally homeomorphic. This is a consequence of the last result presented in this chapter, the Theorem on Invariance of Domain 1.5.8, and proves that the dimension on manifolds is well defined.

1.1 Basic concepts

The objects considered in this chapter are topological manifolds, that means Hausdorff spaces, which are locally homeomorphic to some $\mathbb{R}^n$ and have a countable basis. The integer n is called the dimension of the manifold.

A *chart* (or *coordinate system*) of an n-dimensional manifold X is a homeomorphism $h : U \to U' \subseteq \mathbb{R}^n$ of an open set $U \subseteq X$ (called chart domain) onto an open set $U' \subseteq \mathbb{R}^n$.

A collection of charts $\{h_\alpha\}_{\alpha \in A}$ is an *atlas* for X if the union of the chart

domains U_α covers the manifold, $\bigcup_{\alpha\in A} U_\alpha = X$.

Given two charts $h_\alpha : U_\alpha \to U'_\alpha \subseteq \mathbb{R}^n$ and $h_\beta : U_\beta \to U'_\beta \subseteq \mathbb{R}^n$, they are both defined on the intersection of the domains $U_\alpha \cap U_\beta$. By composition, we obtain a homeomorphism $h_{\alpha\beta} = h_\beta \circ h_\alpha^{-1} : h_\alpha(U_\alpha \cap U_\beta) \to h_\beta(U_\alpha \cap U_\beta)$ of open sets in $\mathbb{R}^n$, which is called chart transformation.

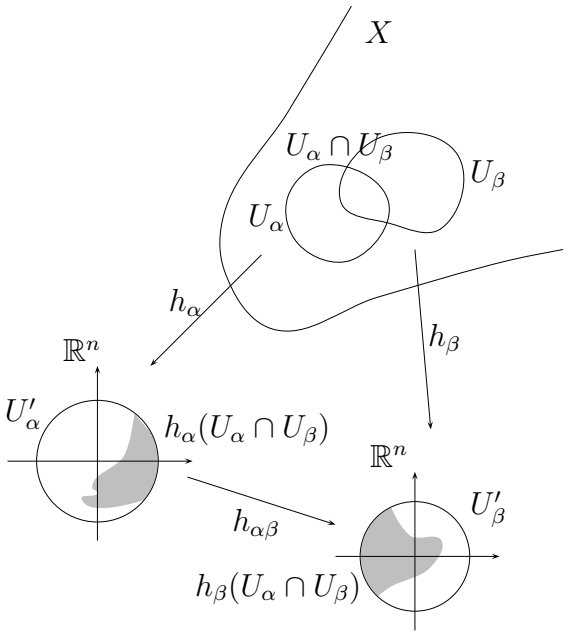

Figure 1.1: Charts and a chart transformation.

An atlas of a manifold is called differentiable of class C^k if all its chart transformation are differentiable of class C^k.

A *C^k-differentiable structure* on a manifold is a maximal C^k-differentiable atlas. A *C^k-differentiable manifold* is a topological manifold together with a C^k-differentiable structure.

A continuous map $f : X \to Y$ between C^k-differentiable manifolds is said to be *C^k-differentiable* at a point $p \in X$ if for some (and therefore for any) chart $h : U \to U'$ and $k : V \to V'$ with $p \in U$, $f(p) \in V$, the composition $k \circ f \circ h^{-1}$ is C^k-differentiable at the point $h(p)$.

A *C^k-diffeomorphism* is an invertible C^k-differentiable map (that means, it is bijective and its inverse map is also C^k-differentiable).

In what follows we will suppose that all manifolds and maps are of class C^∞ and we will refer to them simply as differentiable, or smooth.

Let X be a $(n+k)$-dimensional differentiable manifold. A subset $Y \subseteq X$ is called an *n-dimensional differentiable submanifold of X* if for every point

$p \in Y$ there exists a neighbourhood $U \subseteq X$ and a chart around p

$$h : U \to U' \subseteq \mathbb{R}^{n+k},$$

such that $h(Y \cap U) = U' \cap \mathbb{R}^n$, where we consider $\mathbb{R}^n = \mathbb{R}^n \times \{\underline{0}\} \subseteq \mathbb{R}^n \times \mathbb{R}^k$. The subset Y becomes a manifold, taking as an atlas all the maps $h_{|Y \cap U}$, where h is a chart of X. The number k is called the *codimension* of the submanifold.

A differentiable map $f : X \to Y$ between differentiable manifolds is an *embedding* if $f(X) \subseteq Y$ is a differentiable submanifold, and $f : X \to f(X)$ is a diffeomorphism.

Note that every differentiable structure has a well defined dimension. In fact, there can not exist charts in the same differentiable structure with codomains $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, if $n \neq m$, since there is no map $\mathbb{R}^n \to \mathbb{R}^m$ which is invertible and differentiable.

At this point, we could think that some manifold can have two differentiable structures with different dimensions. We will see, as a consequence of the Theorem on Invariance of Domain 1.5.8, that this is never the case, and therefore the dimension of every manifold is always well defined.

1.2 Tangent space

Given two manifolds X and Y, and a point $p \in X$ with an open neighborhood $U \subseteq X$, we say that two differentiable maps $f, g : U \to Y$ are equivalent if there is an open neighbourhood $V \subseteq U$ of p such that $f_{|V} = g_{|V}$.

An equivalence class for this relation is called a *germ of a map $X \to Y$ at p*. We denote by $\overline{f} : (X, p) \to Y$ or $\overline{f} : (X, p) \to (Y, f(p))$ the germ defined by the differentiable map f.

A *function germ* is a differentiable germ $(X, p) \to \mathbb{R}$. We denote by $\epsilon(p)$ the set of all function germs around $p \in X$. It is a real algebra. We will write ϵ_n for the set of germs $(\mathbb{R}^n, 0) \to \mathbb{R}$.

Given a germ $\overline{f} : (X, p) \to (Y, q)$, the mapping

$$\begin{array}{rcl} f^* : \epsilon(q) & \to & \epsilon(p) \\ \overline{\phi} & \mapsto & \overline{\phi} \circ \overline{f} \end{array}$$

is a contravariant functor between the algebras of all functions germs around the points.

A *derivation on $\epsilon(p)$* is a linear map $\delta : \epsilon(p) \to \mathbb{R}$ which satisfies the product rule

$$\delta(\overline{\phi} \circ \overline{\psi}) = \delta(\overline{\phi}) \circ \overline{\psi}(p) + \overline{\phi}(p) \circ \delta(\overline{\psi}).$$

Definition 1.2.1
The *tangent space* T_pX of a manifold X at a point p is the real vector space of the derivations of $\epsilon(p)$.

If δ is a derivation of $\epsilon(p)$, one can check that $\delta \circ f^*$ is a derivation of $\epsilon(f(p))$, hence we can define the mapping

$$\begin{aligned} T_pf : \ T_pX &\to T_{f(p)}Y \\ \delta &\mapsto \delta \circ f^*, \end{aligned}$$

called the *differential of f at p*, by setting

$$T_pf(\delta)(\overline{\phi}) = \delta \circ f^*(\overline{\phi}) := \delta(\overline{\phi} \circ \overline{f})$$

for each germ $\overline{\phi} : (Y, q) \to \mathbb{R}$.

If $h = (x_1, \ldots, x_n)$ is a chart of X, among the derivations there are the partial derivatives,

$$\begin{aligned} \frac{\partial}{\partial x_i} : \ \epsilon_n &\to \mathbb{R} \\ \overline{\phi} &\mapsto \frac{\partial}{\partial x_i}\phi(0). \end{aligned}$$

The following theorem holds (see [BJ82], Theorem (2.4)).

Theorem 1.2.2
Let X and Y be manifolds of dimension n and m respectively. Introduce local coordinates $(x_1, \ldots, x_n)$ around the point $p \in X$ and $(y_1, \ldots, y_m)$ around $q \in Y$. Then the derivations $\frac{\partial}{\partial x_i}$, $\frac{\partial}{\partial y_j}$ form vector space bases of T_pX and T_qY respectively, and the tangential map of a germ $\overline{f} : (X, p) \to (Y, q)$ with respect to these bases is given by the Jacobi matrix calculated in 0, $Df_0 : \mathbb{R}^n \to \mathbb{R}^m$.

Given a differentiable manifold of dimension n, the tangent space in each of its points is a vector space of dimension n. Once again, we see that the dimension of the manifolds is well defined.

A differentiable germ $\overline{f} : (X, p) \to (Y, q)$ is *invertible* if there is a mapping f which defines the germ $\overline{f}$ and which maps a neighbourhood of p diffeomorphically onto a neighbourhood of q.

For example, given an n-dimensional manifold X and a point $p \in X$, one can find a chart around p which defines an invertible germ $\overline{h} : (X, p) \to (\mathbb{R}^n, 0)$, and hence an isomorphism $h^* : \epsilon_n \to \epsilon(p)$.

The following theorem (see [BJ82], Theorem (5.1)) holds.

Theorem 1.2.3 (Inverse Function Theorem)
A differentiable germ is invertible if and only if its differential is an isomorphism.

The *rank* of a differentiable map $f : X \to Y$ at a point $p \in X$ is

$$\operatorname{rk}_p f = \operatorname{rk} T_p f.$$

This is also the rank of the germ $\overline{f} : (X, p) \to Y$.

The rank is lower semi-continuous, that means, if $\operatorname{rk}_p f = r$, then there is a neighbourhood U of p such that $\operatorname{rk}_q f \geq r$ for all $q \in U$.

A germ $\overline{f} : (X, p) \to (Y, q)$ is said to have constant rank r if there is a function f in the germ such that in some open neighbourhood of p the rank of f is constant.

Theorem 1.2.4 (Rank Theorem)
Let $\overline{f} : (X, p) \to (Y, q)$ be a germ of constant rank r, where X has dimension n and Y has dimension m. Then, there are charts h around p and k around q such that the germ

$$\overline{k} \circ \overline{f} \circ \overline{h}^{-1} : (\mathbb{R}^n, 0) \to (\mathbb{R}^m, 0)$$

is represented by the map

$$(x_1, x_2, \ldots, x_n) \mapsto (x_1, x_2, \ldots, x_r, 0, \ldots, 0).$$

For the proof, we refer to [BJ82], Theorem (5.4).

Definition 1.2.5
Let $f : X \to Y$ be a C^1 map.
A point $p \in X$ is *regular* or *smooth* if the differential $T_p f$ is surjective.
A point $q \in Y$ is a *regular value* if all points of its preimage $f^{-1}(q)$ are regular.
A point is *critical* or *singular* if it is non-regular.
A differentiable map $f : X \to Y$ is called a *submersion* if $\operatorname{rk}_p f = \dim Y$ for all $p \in X$, that is the same as to require that every point of X is regular, or that every point of $f(X)$ is a regular value.
The map f is called an *immersion* if $\operatorname{rk}_p f = \dim X$ for all $p \in X$, that is the differential $T_p f$ is injective at every point of X.

It is not hard to see that from the Rank Theorem, we can derive the following theorem.

Theorem 1.2.6
If $f : X \to Y$ is an injective immersion and $f : X \to f(X)$ a homeomorphism, where $f(X) \subseteq Y$ has the subspace topology, then f is an embedding.

Proposition 1.2.7
Let q be a regular value of the differentiable map $f : X \to Y$, where $\dim X > \dim Y = n$. Then $f^{-1}(q)$ is either empty or a differentiable submanifold of X with codimension n.

For the proof see [Milnor78] §2, Lemma 1.

1.3 Sard's Theorem

We will state and prove the Sard's Theorem 1.3.5, which ensures that the set of critical values of a differentiable mapping of manifolds is "small" in some sense that we will see.

Definition 1.3.1
A subset $A \subseteq \mathbb{R}^n$ has Lebesgue measure zero if for every $\varepsilon > 0$ it can be covered by countably many n-cubes of edge w with total n-volume $\sum_{i \in \mathbb{N}} w^n < \varepsilon$.

Remark 1.3.2
In the definition it does not matter if the cubes are open or closed, and we may also take n-spheres instead of n-cubes.
A countable union of sets of measure zero has again measure zero.

Proposition 1.3.3
Let U be an open subset of $\mathbb{R}^n$ and $A \subseteq U$ a subset of measure zero. If $f : U \to \mathbb{R}^m$ is differentiable, then $f(A)$ has measure zero.

Proof. Since U is open, it is the union of a countable sequence of compact balls, and we may assume that A is contained in a compact ball, and that the cubes of a covering of A are also contained in a larger compact ball $K \subseteq U$.

By the mean value theorem, for $[x, x + h] \in K$ there is a constant c such that

$$| f(x + h) - f(x) | \leq c | h |$$

If a cube $W \subseteq K$ has edge length a, we have that $| x - x' | \leq a\sqrt{n}$ for $x \in W$, and $| f(x) - f(x') | \leq ca\sqrt{n}$. Hence, $f(W)$ lies in a cube of volume $(2c\sqrt{n})^m | W |$ and, since the constant $(2c\sqrt{n})^m$ is independent of the cube, the assertion follows. □

This proposition allows us to speak of subsets with measure zero in a differentiable manifold.

Definition 1.3.4
A subset A of a differentiable manifold has measure zero if for every chart $h : U \to \mathbb{R}^n$ the image $h(A \cap U) \subseteq \mathbb{R}^n$ has measure zero.

Theorem 1.3.5 (Sard's Theorem)
The set of critical values of a differentiable mapping of manifolds has Lebesgue measure zero.

In particular, by Proposition 1.2.7, if $f : X \to \mathbb{R}^n$ is differentiable, then the preimage of a point is either empty or it is an n-codimensional submanifold for almost all points of X.

Since a manifold has a countable basis for its topology, each atlas has a subatlas with countable many charts. Applying Proposition 1.3.3 to the chart transformations, we obtain that A has measure zero if and only if for all charts h_α of a fixed atlas, the image $h_\alpha(A \cap U_\alpha) \subseteq \mathbb{R}^n$ has measure zero.

Thanks to this result, in order to prove Sard's Theorem, it is only necessary to consider subsets of $\mathbb{R}^n$.

We also need another result, Fubini's Theorem.

Theorem 1.3.6 (Fubini)
Define

$$\mathbb{R}^{n-1}_t = \{x \in \mathbb{R}^n \mid x_n = t\},$$

let A be a compact subset of $\mathbb{R}^n$ and suppose $A_t = A \cap \mathbb{R}^{n-1}_t$ has measure zero in $\mathbb{R}^{n-1}_t \cong \mathbb{R}^{n-1}$ for all $t \in \mathbb{R}$. Then A has measure zero in $\mathbb{R}^n$.

Proof. Since A is compact, we may assume that $A \subseteq \mathbb{R}^{n-1} \times I$, where $I = [0, 1]$, and that A_t has measure zero in $\mathbb{R}^{n-1} \times \{t\}$. Thus, for any $\varepsilon > 0$ there is a countable covering of A_t, whose elements are open cubes $W^i_t \subseteq \mathbb{R}^{n-1}_t$, such that

$$\sum_{i \in \mathbb{N}} |\, W^i_t \,| < \varepsilon$$

Let W_t be the projection of $\cup_i W^i_t$ on the first factor $\mathbb{R}^{n-1}$ of $\mathbb{R}^{n-1} \times I$. For a fixed t, the function $|\, x_n - t \,|$ is continuous on A, and vanishes exactly on A_t. Furthermore, outside of $W_t \times I$, this function has a positive minumum value α, because A is compact and W_t is open.

Letting I_t be the interval $(t - \alpha, t + \alpha)$, we have

$$\{x \in A \mid |\, x_n - t \,| < \alpha\} \subseteq W_t \times I_t$$

The intervals I_t cover I, hence there is a finite subcovering $\{I_j\}_{j=1,\dots,k}$ with $I_j = I_{t_j}$ for some $t_j \in I$. We may assume that their total volume $\sum \mid I_j \mid$ does not exceed 2, because we can chose the finite subcovering in such a manner that each point belongs to at most two open sets of the subcovering: if a point lies in more than two open sets, then there is one of them with a smallest initial point and one with a largest end point. We can cancel from the covering all other sets containing that point.

The sets $\{W^i_{t_j} \times I_j \mid j = 1, \dots, k; i \in \mathbb{N}\}$ have total volume not exceeding 2ε and our set A is contained in their union. Hence, A has measure zero in $\mathbb{R}^n$. □

It is not necessary that the subset A is compact, it is sufficient that it is a countable union of compact sets. Then, Fubini's Theorem holds for closed sets, open sets, images of open or closed sets under continuous maps, countable unions and finite intersections of open or closed sets.

Proof of Sard's Theorem 1.3.5. We have to show the following:
Let $U \subseteq R^n$ be open, $f : U \to \mathbb{R}^m$ differentiable and $C \subseteq U$ the set of critical points of f. Then $f(C) \subseteq \mathbb{R}^m$ has measure zero.

We proceed by induction on n. If $n = 0$, then R^n is a point, and $f(C)$ is at most one point, and the theorem holds.

Suppose now the result is true for $n - 1$; we show it for n. Let $C_i \subseteq U$ be the set of points at which all partial derivatives of order up to i vanish. Then the sets C_i form a decreasing sequence of closed sets

$$C \supseteq C_1 \supseteq C_2 \supseteq \dots.$$

Note that C is a disjoint union

$$C = (C \smallsetminus C_1) \sqcup (C_1 \smallsetminus C_2) \sqcup \dots \sqcup (C_{k-1} \smallsetminus C_k) \sqcup C_k.$$

We will show that

(*a*) $f(C \smallsetminus C_1)$ has measure zero

(*b*) $f(C_i \smallsetminus C_{i+1})$ has measure zero for $i \geq 1$

(*c*) there is a $\overline{k}$ such that $f(C_k)$ has measure zero for all $k \geq \overline{k}$.

Then, C is a countable union of sets of measure zero, hence we are done.

Note that all these sets are in the class to which Fubini's Theorem is applicable. Furthermore, it is sufficient to show that each point of the set D

we are considering ($D = C \setminus C_1$ or $C_i \setminus C_{i+1}$ or C_k) has a neighbourhood V such that $f(V \cap D)$ has measure zero, since each set D is covered by countably many such neighbourhoods, because $\mathbb{R}^n$ is paracompact.

Now we prove (a).

If $m = 1$, then $C = C_1$ and the claim is trivial. Hence, assume $m \geq 2$. Let $x \in C \setminus C_1$, then some partial derivatives do not vanish at the point x. Assume $\frac{\partial f_1}{\partial x_1}(x) \neq 0$. Then the map

$$h : U \to \mathbb{R}^n \qquad (x_1, \ldots, x_n) \mapsto (f_1(x_1), x_2, \ldots, x_n)$$

is not singular at the point x. Therefore, the restriction of this map to a neighbourhood V of x is a chart $h_{|V} = h : V \to V'$ and the transformed map $g = f \circ h^{-1} : V' \to \mathbb{R}^m$ has the form

$$g(z_1, \ldots, z_n) = (z_1, g_2(z), \ldots, g_m(z))$$

locally about $h(x)$.

For a fixed $t \in \mathbb{R}$, this map takes the hyperplane $\{z \mid z_1 = t\}$ into the plane $\{y \mid y_1 = t\}$. Restrict g to

$$g^t : (\{t\} \times \mathbb{R}^{n-1}) \cap V' \to \{t\} \times \mathbb{R}^{m-1}$$

Then a point in $\{t\} \times \mathbb{R}^{n-1}$ is critical for g if and only if it is critical for g^t, because the Jacobi matrix of g is

$$\left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline * & & & \\ \vdots & & Dg^t & \\ * & & & \end{array}\right)$$

By the induction hypothesis, the set of critical values of g^t has measure zero in $\{t\} \times \mathbb{R}^{n-1}$, hence the set of critical values of g has an intersection of measure zero with each hyperplane $\{y \mid y_1 = t\}$. Thus, by Fubini's Theorem, also the set of critical values of g has measure zero in $\mathbb{R}^n$, and (a) holds.

To prove (b), we proceed in the same way. For each point $x \in C_k \setminus C_{k+1}$ there is some derivative of order $(k+1)$, which does not vanish at the point x. We may assume that

$$\frac{\partial^{k+1} f_1}{\partial x_1 \partial x_{\nu_1} \ldots \partial x_{\nu_k}}(x) \neq 0$$

Define the function $w : U \to \mathbb{R}$, $w = \frac{\partial^k f_1}{\partial x_{\nu_1} \ldots \partial x_{\nu_k}}$. Then $w(x) = 0$ since $x \in C_k$, and $\frac{\partial w}{\partial x_1}(x) \neq 0$ since $x \notin C_{k+1}$. The map

$$h : x \mapsto (w(x), x_2, \ldots, x_n)$$

defines a chart $h : V \to V'$ about x, and we have

$$h(C_k \cap V) \subseteq \{0\} \times \mathbb{R}^{n-1} \subseteq \mathbb{R}^n$$

Consider the transformed map $g = f \circ h^{-1} : V' \to \mathbb{R}^m$. By the induction hypothesis, its restriction

$$g^0 : (\{0\} \times \mathbb{R}^{n-1}) \cap V' \to \mathbb{R}^m$$

has a set of critical values of measure zero in $\mathbb{R}^m$.

Each point from $h(C_k \cap V)$ is critical for g^0 because all partial derivatives of g of order up to k vanish, and hence the same do the derivatives of g^0. Thus, $f(C_k \cap V) = g \circ h(C_k \cap V)$ has measure zero.

To prove (c), suppose $W \subseteq U$ be a cube with edges of length a and let $k > \frac{n}{m} - 1$. We show that $f(W \cap C_k)$ has measure zero. This will be sufficient, since U is a countable union of cubes.

For $x \in C_k \cap W$ and $x + h \in W$, by the Taylor formula we have

$$f(x+h) = f(x) + \underbrace{f^{(1)}(x) + \ldots + f^{(k)}(x)}_{0} + R(x,h),$$

with

$$|\, R(x,h) \,| \leq \frac{f^{(k+1)}(x_1)}{(k+1)!} \,|\, h \,|^{\,k+1} = c \,|\, h \,|^{\,k+1},$$

where c is a fixed constant depending only on f and W. Decompose W into r^n cubes with edges of length $\frac{a}{r}$ and suppose W_1 is a cube of the decomposition, which contains a point $x \in C_k$. Then each point from W_1 can be written as $x + h$, where

$$|\, h \,| \leq \frac{a}{r}\sqrt{n}$$

since this is the length of the diagonal of the n-cube W_1. From the remainder estimate of the Taylor formula, $f(W_1)$ lies in a cube with edges of length

$$2c \,|\, h \,|^{\,k+1} \leq 2c \left(\frac{a}{r}\sqrt{n}\right)^{k+1} = \frac{b}{r^{k+1}}$$

where b is a constant dependng only on f and W and not on the decomposition. Alle these cubes together have a total volume not exceeding

$$r^n \frac{b^m}{r^{m(k+1)}}$$

and for $m(k+1) > n$ this volume converges to zero as r increases. Hence, chosing a decomposition into sufficiently small cubes, the total volume can be made arbitrarily small. □

Corollary 1.3.7
The regular values of a differentiable map $f : X \to Y$ are dense in Y.

We prove in Corollary 1.3.10 that spheres with different dimensions can not be homeomorphic. To do that, we use the Sard's Theorem and the following result.

Corollary 1.3.8
Let $f : X \to Y$ be a differentiable map between manifolds of dimension n and m respectively, with $n < m$. Then $f(X)$ has measure 0.

This follows directly from the fact that all points in $f(X)$ are critical values.

We recall the definitions of homotopy and smooth homotopy. Two maps $f, g : X \to Y$ are called homotopic if there is a map $F : X \times I \to Y$ such that

$$F(x,0) = f(x) \qquad F(x,1) = g(x)$$

for all $x \in X$. The map F is called a homotopy between f and g. If in addition the homotopy F is differentiable, then it is called a smooth homotopy and f and g are said to be smoothly homotopic.

A mapping is called (smoothly) homotopically trivial if it is (smoothly) homotpic with a constant map.

Theorem 1.3.9
Let $n < m$ and

$$f : S^n \to S^m \subseteq \mathbb{R}^{m+1}$$

be a continuous function. Then f is homotopically trivial.

Proof. If f is differentiable, then it is not surjective by Corollary 1.3.8 and hence homotopically trivial.

We can extend f to an open subset U of $\mathbb{R}^{n+1}$ with a continuous function $\tilde{f} : U \to S^m$. Then $\tilde{f}_{|S^n} = f$.

Let $K \subseteq U$ be a compact neighbourhood of S^n. Using an approximation theorem (see for example [Hir76], chapter 2), we can approximate $\tilde{f}$ on K with a polynomial function

$$g : K \to \mathbb{R}^{m+1}.$$

Then $g_{|S^n} : S^n \to \mathbb{R}^{m+1}$ approximates f and since $|\, f(x)\,| = 1$, we may assume $0 \notin g(S^n)$.

Let $\hat{g} : S^n \to S^m$ be defined through $\hat{g}(x) = \frac{g(x)}{|\, g(x)\,|}$. It is well defined since $0 \notin g(S^n)$ and it is differentiable, since g is a polynomial function. Define the map

$$\begin{array}{rccc} F: & S^n \times I & \to & S^m \\ & (x,t) & \mapsto & \dfrac{(1-t)f(x)+tg(x)}{|\,(1-t)f(x)+tg(x)\,|}. \end{array}$$

We claim that this is a homotopy between f and $\hat{g}$. In fact, $F(x,0) = f(x)$ and $F(x,1) = \frac{g(x)}{|\, g(x)\,|} = \hat{g}(x)$ for each $x \in S^n$ and we show that the denominator is not 0. Since $g_{|S^n}$ approximates f, we can assume $|\, f(x) - g(x)\,| < 1$ for all $x \in S^n$. Furthermore, we know that $|\, g(x)\,| \neq 0$ and $|\, f(x)\,| = 1$ for all $x \in S^n$. By contradiction, suppose the denominator is 0. Then $t \neq 0$ and $f(x) = t(f(x) - g(x))$, hence

$$1 = |\, f(x)\,| = t\,|\, f(x) - g(x)\,| < t \leq 1,$$

which is a contradiction. Hence $|\,(1-t)f(x) + tg(x)\,| \neq 0$ and F is a homotopy between f and $\hat{g}$.

Since the map $\hat{g} : S^n \to S^m$ is differentiable, it is not surjective, and hence homotopically trivial. Moreover, $\hat{g}$ is homotopic with f, hence f is also homotopically trivial. □

Corollary 1.3.10
If $n < m$, then S^n and S^m are not homeomorphic.

Proof. Suppose by contradiction that $f : S^n \to S^m$ is a homeomorphism, then by Theorem 1.3.9, f is homotopically trivial and so is $f^{-1} \circ f = \mathrm{id}_{S^n}$. Hence, to find a contradiction, it suffices to show that the identity function id_{S^n} is not homotopically trivial.

Let $f : S^n \to S^n$ be differentiable, and q a regular value. Then by Proposition 1.2.7, $f^{-1}(q)$ is a submanifold of dimension 0, i.e. $f^{-1}(q) = \{p_1, \ldots, p_k\}$.
The differential of f at a point p_i is a map $T_{p_i}f : T_{p_i}S^n \to T_qS^n$. Choose an orientation for T_qS^n and define

$$\varepsilon(p_i) = \begin{cases} +1 & \text{if } T_{p_i}S^n \text{ is consistently oriented with } T_qS^n \\ -1 & \text{otherwise.} \end{cases}$$

It can be proved that the number

$$d_q(f) = \sum_{i=1}^{k} \varepsilon(p_i)$$

does not depend on the choice of the regular value q. We call this number the degree of f, $\deg f$. One can also prove that if f and g are smoothly homotopic, then $\deg f = \deg g$. Both proofs can be found in [Milnor78], §5. Furthermore, it is easy to see that $\deg \mathrm{id} = 1$ and $\deg c = 0$ for each constant map c. That means the identity map is not smoothly homotopically trivial.

To complete the proof, it can be shown that if id_{S^n} were homotopically trivial, then it would also be smoothly homotopically trivial.The proof uses approximation on compact sets and is quite similar to the first part of the proof of Theorem 1.3.9. □

1.4 Some results in Algebraic Topology

Many results stated so far in this chapter can be obtained more directly if we use the homology theory: that is what we do in the rest of the chapter.

We recall some results without giving the proofs, which can be found in the book of Spanier [Spa66] or Greenberg [Gre81] for example.

Proposition 1.4.1
A space X consisting in only one point is acyclic, that is $H_0(X) \cong \mathbb{Z}$ and $H_k(X) = (0)$ for $k \geq 1$.

Proposition 1.4.2
Let X be a topological space and $X = \dot{\cup}_{i\in I} X_i$ its decomposition in path connected components. Then for each $k \geq 0$ holds:

$$H_k(X) = \bigoplus_{i\in I} H_k(X_i).$$

Proposition 1.4.3
If X is a nonempty path connected space, then $H_0(X) \cong \mathbb{Z}$.

Theorem 1.4.4 (Homotopy Theorem)
Let $f, g : X \to Y$ be hotomopic maps between topological spaces. Then the induced homomorphisms $f_*, g_* : H_k(X) \to H_k(Y)$ coincide for each $k \geq 0$.

Theorem 1.4.5 (Mayer-Vietoris)
Let U and V be open subsets of a topological space X such that $X = U \cup V$. Then, there is a long homology sequence

$$\ldots \xrightarrow{\nu} H_{k+1}(X) \xrightarrow{\Delta} H_k(U \cap V) \xrightarrow{\mu} H_k(U) \oplus H_k(V) \xrightarrow{\nu} H_k(X) \xrightarrow{\Delta} \ldots$$

Theorem 1.4.6 (Homology group of the sphere)
For every $n \geq 1$ holds:

$$H_k(S^n) = \begin{cases} \mathbb{Z} & \text{if } k = 0 \text{ or } k = n \\ (0) & \text{otherwise.} \end{cases}$$

Theorem 1.4.7
$\mathbb{R}^n$ and $\mathbb{R}^m$ are homeomorphic if and only if $n = m$.

Proof. Assume there exists a homeomorpism $f : \mathbb{R}^n \to \mathbb{R}^m$. Let $p \in \mathbb{R}^n$ such that $f(p) = 0 \in \mathbb{R}^m$. Then $f_{|\mathbb{R}^n \setminus \{p\}} : \mathbb{R}^n \setminus \{p\} \to \mathbb{R}^m \setminus \{0\}$ is also a homeomorphism. Since $\mathbb{R}^n \setminus \{p\}$ has the same homotopy type as S^{n-1}, by the Homotopy Theorem 1.4.4 we have

$$H_{m-1}(S^{n-1}) \cong H_{m-1}(\mathbb{R}^n \setminus \{p\}) \xrightarrow[\cong]{f_*} H_{m-1}(\mathbb{R}^m \setminus \{0\}) \cong H_{m-1}(S^{m-1})$$

If $m = 1$, since $H_0(S^{n-1}) \cong H_0(S^0) \cong \mathbb{Z} \oplus \mathbb{Z}$, it follows that $n = 1$.

If $m \geq 2$, since $H_{m-1}(S^{n-1}) \cong H_{m-1}(S^{m-1}) \cong \mathbb{Z}$, by virtue of Theorem 1.4.6, it follows that $n = m$. □

1.5 Invariance of domain

Definition 1.5.1
A continuous function $f : X \to Y$ is called an *embedding* if $X \cong f(X)$, where $f(X) \subseteq Y$ has the topology induced by the topology of Y.
Let $\mathbb{D}^r$ be a r-disc and $f : \mathbb{D}^r \to Y$ an embedding. Then $f(\mathbb{D}^r)$ is called a (r-)*ball* in Y. Note that to define a ball we can also use I^r instead of $\mathbb{D}^r$, since they are homeomorphic.
Let $f : S^r \to Y$ be an embedding. Then $f(S^r)$ is called a (r-)*sphere* in Y.

Proposition 1.5.2
Let $\{U\}_{j \in J}$ be an open covering of a compact metric space X. Then there is a Lebesgue number $\varepsilon > 0$ such that for each subset $A \subseteq X$ with diam $A < \varepsilon$ *there is an index $j \in J$ such that $A \subseteq U_j$.*

Lemma 1.5.3
Let $B \subset S^n$ be a r-ball ($r < n$); then $S^n \setminus B$ is acyclic, that is $H_0(X) \cong \mathbb{Z}$ and $H_k(X) = (0)$ for $k \geq 1$.

Proof. Let $B = f(\mathbb{D}^r) \subset S^n$, where $f : \mathbb{D}^r \to S^n$ is an embedding. We proceed by induction on the dimension r of the ball.

If $r = 0$, $\mathbb{D}^0$ is a point, hence $S^n \setminus B \cong S^n \setminus \{point\} \cong \mathbb{R}^n$, that is acyclic.

Assume the result holds for $r-1$. Let $X = S^n \setminus B$, and $z \in Z_k(X)$ a cycle for $k \geq 1$, $z = y - x$ for $k = 0$, where x and y are points of X. We want to prove that z is a boundary, i.e. there is a chain $c \in S_{k+1}(X)$ such that $\partial c = z$. (Then $H_0(X) \cong \mathbb{Z}$ and for $k \geq 1$ we have $[z] = 0$, also $H_k(X) = (0)$).

Since $\mathbb{D}^r \cong I^r$, we can assume that $B = f(I^r)$, where $f : I^r \to B \subset S^n$ is a homeomorphism. For each $t \in I$ define $B_t = f(I^{r-1} \times \{t\})$; it is a $(r-1)$-ball in S^n.

Let $X_t = S^n \setminus B_t$; by the induction hypothesis there is a $c_t \in S_{k+1}(X_t)$ such that $\partial c_t = z$.

The members of c_t have compact images and B_t is compact, hence there is an open neighbourhood $U_t \subseteq S^n$ of B_t such that $c_t \in S_{k+1}(X \setminus U_t)$. Let $V_t = f^{-1}(U_t) \subseteq I^r$. Since I^{r-1} is compact and $I^{r-1} \times \{t\} \subseteq V_t$, there is an open neighbourhood $I_t \subseteq I$ of t, such that $I^{r-1} \times I_t \subseteq V_t$.

Let ε be a Lebesgue number for the open covering $\{I_t\}_{t \in I}$ of I and choose an integer m such that $m \geq \frac{1}{\varepsilon}$. For $j = 1, \ldots, m$ there is a $t_j \in I$ such that $I_j = \left[\frac{j-1}{m}, \frac{j}{m}\right] \subseteq I_{t_j}$. Let $Q_j = f(I^{r-1} \times I_j)$ and $c_j = c_{t_j}$ for $j = 1, \ldots, m$. Then for each j: $c_j \in S_{k-1}(S^n \setminus Q_j)$ and $\partial c_j = z$.

We have $B = \bigcup_{j=1}^m Q_j$, and we may assume $m \geq 2$. For each j let $X_j = S^n \setminus Q_j$.

The couple (X_1, X_2) is excissive for $X_1 \cup X_2 = S^n \setminus (Q_1 \cap Q_2) \cong S^n \setminus I^{r-1}$, hence by the Mayer-Vietoris Theorem the following sequence is exact:

$$\begin{array}{rcl} H_{k+1}(X_1 \cup X_2) \;\to\; H_k(X_1 \cap X_2) & \xrightarrow{\mu} & H_k(X_1) \oplus H_k(X_2) \\ & [z]_{X_1 \cap X_2} \;\mapsto & ([z]_{X_1}, -[z]_{X_2}) \end{array}$$

By the induction hypothesis $X_1 \cup X_2$ is acyclic, that is $H_{k+1}(X_1 \cup X_2) = (0)$, thus the kernel of μ is (0), that means μ is injective.

Since $[z]_{X_1} = [\partial c_1]_{X_1} = 0$ and $[z]_{X_2} = [\partial c_2]_{X_2} = 0$, there is a $c_{12} \in S_{k+1}(X_1 \cap X_2) = S_{k+1}(S^n \setminus (Q_1 \cup Q_2))$ such that $\partial c_{12} = z$, because $[z]_{X_1 \cap X_2} = 0$.

Repeat this argument with $\tilde{X}_1 = S^n \setminus (Q_1 \cup Q_2)$ and $X_3 = S^n \setminus Q_3$. Then there is a $c_{123} \in S_{k+1}(S^n \setminus (Q_1 \cup Q_2 \cup Q_3))$ with $\partial c_{123} = z$.

After $m-1$ times we get a $c = c_{12\ldots m} \in S_{k+1}(S^n \setminus (Q_1 \cup \cdots \cup Q_m)) = S_{k+1}(S^n \setminus B)$ such that $\partial c = z$. □

Theorem 1.5.4
Let $n \geq 1$, $0 \leq r \leq n-1$ and $S \subset S^n$ an r-sphere.
Then for $0 \leq r \leq n-2$ holds:

$$H_k(S^n \setminus S) \cong \begin{cases} \mathbb{Z} & \text{for } k = 0 \text{ and } k = n-r-1 \\ (0) & \text{otherwise.} \end{cases}$$

For $r = n - 1$ holds:

$$H_k(S^n \setminus S) \cong \begin{cases} \mathbb{Z} \oplus \mathbb{Z} & \text{for } k = 0 \\ (0) & \text{otherwise.} \end{cases}$$

Proof. We proceed by induction on r.

Let $r = 0$. S^0 is the union of two distinct points, $S^n \setminus S^0 = S^n \setminus \{p, q\}$. Since $S^n \setminus \{p\} \cong \mathbb{R}^n$, it follows that $S^n \setminus S^0 \cong \mathbb{R}^n \setminus \{0\}$ which is homotopically equivalent to S^{n-1}, since S^{n-1} is a deformation retract of $\mathbb{R}^n \setminus \{0\}$. The claim follows from Theorem 1.4.6.

Now, supposing the result is true for $r - 1$, we prove it for r. Let $D^+ = \{z \in S^r \mid z_{r+1} \geq 0\}$ and $D^- = \{z \in S^r \mid z_{r+1} \leq 0\}$. Let $f : S^r \to S \subseteq S^n$ be an embedding and $B^+ = f(D^+)$, $B^- = f(D^-) \subseteq S$. Then $S = B^+ \cup B^-$; furthermore $T = B^+ \cap B^- = f(D^+ \cap D^-)$ is a $(r-1)$-sphere in S^n.

Let $U = S^n \setminus B^+$, $V = S^n \setminus B^-$; then $U \cup V = S^n \setminus T$ and $U \cap V = S^n \setminus S$.

For $k \geq 1$ we have the Mayer-Vietoris sequence

$$\begin{aligned}(0) = H_{k+1}(U) \oplus H_{k+1}(V) \;\to\; & H_{k+1}(S^n \setminus T) \\ & \xrightarrow[\cong]{\Delta} H_k(S^n \setminus S) \;\to\; H_k(U) \oplus H_k(V) = (0)\end{aligned}$$

because U and V are acyclic by Lemma 1.5.3.
By the induction hypothesis,

$$H_k(S^n \setminus S) \cong H_{k+1}(S^n \setminus T) \cong \begin{cases} \mathbb{Z} & \text{for } k = n - r - 1 \\ (0) & \text{for } k \geq 1,\ k \neq n - r - 1 \end{cases}$$

For $k = 0$ we have:

$$\begin{aligned}\mathbb{Z} \oplus \mathbb{Z} \cong H_0(U) \oplus H_0(V) &\xrightarrow{\nu} H_0(U \cup V) = H_0(S^n \setminus T) \cong \mathbb{Z} \\ ([x], [y]) &\mapsto [x] + [y]\end{aligned}$$

Then $\ker \nu \cong \langle [x_0], -[x_0] \rangle \cong \mathbb{Z}$ for $x_0 \in S^n \setminus S$. Hence, the sequence

$$(0) = H_1(U) \oplus H_1(V) \to H_1(S^n \setminus T) \xrightarrow{\Delta} H_0(S^n \setminus S) \xrightarrow{\mu} \ker \nu \cong \mathbb{Z} \to (0)$$

is exact and splits because $\mathbb{Z}$ is a free group. Then $H_0(S^n \setminus S) \cong H_1(S^n \setminus T) \oplus \mathbb{Z}$.

For $r < n - 1$ the induction hypothesis implies $H_1(S^n \setminus T) = (0)$ because $n - r \neq 1$. Then $H_0(S^n \setminus S) \cong \mathbb{Z}$.

For $r = n - 1$ it holds $H_1(S^n \setminus T) = \mathbb{Z}$. Then $H_0(S^n \setminus S) \cong \mathbb{Z} \oplus \mathbb{Z}$. □

Corollary 1.5.5 (Jordan-Brouwer Separation Theorem)
An $(n-1)$-sphere $S \subset S^n$ separates S^n into two components of which it is the common boundary.

Proof. $H_0(S^n \setminus S) \cong \mathbb{Z} \oplus \mathbb{Z}$ by Theorem 1.5.4, hence $S^n \setminus S$ has two path connected components (this follows from Propositions 1.4.2 and 1.4.3). Since $S^n \setminus S$ is an open subset of S^n, it is locally path connected. Let U and V be its path components; they are also its components by Proposition 1.5.6.

U and V are open, hence their boundary is contained in S. We have to show that $S \subseteq \overline{U} \cap \overline{V}$. To prove that, let $x \in S$ and $N \subseteq S^n$ be a neighbourhood of x. Let A be a subset of $S \cap N$ such that $S \setminus A \cong I^{n-1}$. By Lemma 1.5.3, $S^n \setminus (S \setminus A)$ is acyclic, hence it is path connected.

Let $p \in U$ and $q \in V$; then there is a path ω from p to q contained in $S^n \setminus (S \setminus A)$. Since p and q are in different path components of $S^n \setminus S$, ω meets A. Hence A meets $\overline{U}$ and $\overline{V}$. Therefore N meets $\overline{U}$ and $\overline{V}$, and $x \in \overline{U} \cap \overline{V}$ because N was arbitrary. □

Proposition 1.5.6
Let A be a locally path connected space and $A = B \cup C$ its decomposition in path connected components. Then A is not connected.

Proof. We show that if a space is connected and locally path connected, then it is also path connected. Then B and C are obviously connected but $B \cup C$ is not connected, otherwise it would be path connected.

Let D be connected and locally path connected, x a point of D. Let $U = \cup\{V \mid x \in V, V \text{ open and path connected}\}$. It is obviously open and path connected.

Let $y \in \partial U$, then since D is locally path connected, there is an open path connected set W which contains x. Since $y \in \partial U$, $U \cup W$ is not empty, hence $U \cup W$ is open and path connected, hence $U \cup W \subseteq U$, and $y \in U$, which means that U is open and closed. Hence it must be $U = D$ because D is connected. □

Remark 1.5.7
The separation theorem holds also in $\mathbb{R}^n$ for $n \geq 2$: an $(n-1)$-sphere $S \subset \mathbb{R}^n$ separates the euclidean space into two components.

Let $N = (0, \ldots, 0, 1) \in S^n$ be the northpole and $\pi : S^n \setminus \{N\} \to \mathbb{R}^n$ the stereographic projection (which is an homeomorphism). Let $f : S^{n-1} \to \mathbb{R}^n$ be an embedding and $S = f(S^{n-1})$.

Then $M = \pi^{-1}(S) \subseteq S^n$ is an $(n-1)$-sphere in S^n and thus by Theorem 1.5.4, we have $H_0(S^n \setminus M) \cong \mathbb{Z} \oplus \mathbb{Z}$. Hence, $S^n \setminus M = U_1 \dot{\cup} U_2$ has two path

components. We may assume $N \in U_2$. Since $S^n \setminus M$ is open, also U_1 and U_2 are open in S^n.

The component U_2 is path connected, and so is $U_2 \setminus \{N\}$. So $V_1 = \pi(U_1)$ and $V_2 = \pi(U_2 \setminus \{N\})$ are also path connected. $\mathbb{R}^n \setminus S = V_1 \dot{\cup} V_2$ and there is no path from V_1 to V_2.

Theorem 1.5.8 (Theorem on Invariance of Domain)
If U and V are homeomorphic subsets of S^n and U is open in S^n, then V is open in S^n.

Proof. Let $f : U \to V$ be a homeomorphism and let $f(x) = y$. Let B be a neighbourhood of x in U that is homeomorphic to I^n, with boundary S homeomorphic to S^{n-1}. By Lemma 1.5.3, $S^n \setminus f(B)$ is connected, and by Theorem 1.5.5, $S^n \setminus f(S)$ has two components. Since $S^n \setminus f(B)$ and $f(B) \setminus f(S)$ are connected and $S^n \setminus f(S) = (S^n \setminus f(B)) \cup (f(B) \setminus f(S))$, they are the components of $S^n \setminus f(S)$. Hence $f(B) \setminus f(S)$ is an open subset of S^n. Therefore, V is open in S^n, because y was an arbitrary point of $f(B) \setminus f(S) \subseteq V$. □

The Theorem on Invariance of Domain assures that the dimension of a manifold is well defined.

If a point of the manifold has two neighbourhoods U, V, which are homeomorphic to open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, then the intersection $U \cap V$ is also homeomorphic to open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$. Hence, by the Theorem on Invariance of Domain, they are both open in $\mathbb{R}^n$ (supposing $n \geq m$), that means it must be $n = m$, i.e. the dimension is well defined.

Chapter 2

Dimension of affine varieties

There are some different definition of dimension over algebraic varieties, but at the end they all lead to the same result, therefore we could consider them as different characterizations of the dimension property, rather than different definitions.

The most natural one, following the same approach as differential topology, is to consider the tangent space and define the dimension of the variety as the dimension of the tangent vector space. There are at least two different characterizations of the tangent space: one uses the gradient and the Jacobi matrix, the other approach is based on the derivations on the algebra of regular functions.

The tangent space is a local property of the varieties, as it is defined pointwise, and the definition of dimension for the whole variety is then the minimum of the dimensions of the tangent spaces in each point, seen as vector spaces.

An important concept is then that of regular points, whose tangent space has constant dimension in a neighbourhood of the point. For algebraic varieties over $\mathbb{C}$, it turns out that regular points are smooth in the sense of differential topology, that is the variety is a smooth manifold in a neighbourhood of a regular point, and the dimension of the variety coincides with the dimension of the manifold. In other words, the concept of dimension of algebraic varieties and the concept of dimension of complex manifolds coincide on objects belonging to both categories.

In addition to these concepts, there are two definition of dimension, which are purely algebraic: one uses the transcendence degree of the field of rational functions, and the other, called Krull dimension, considers the length of the possible sequences of closed irreducible subsets of a variety, or, equivalently, of the prime ideals in the stalk in a point of the sheaf of regular functions.

2.1 Basic concepts

Throughout this chapter, we denote by $\mathbb{K}$ an algebraically closed field. In this section, we recall some basic definitions and results.

Let $L \subseteq \mathbb{K}[T_1, \ldots, T_n]$ be a set of polynomials in n variables; define the *vanishing set* (set of zeroes) of L as the subset

$$V(L) = \{x \in \mathbb{K}^n \mid f(x) = 0 \text{ for all } f \in L\} \subseteq \mathbb{K}^n.$$

A set $X \subseteq \mathbb{K}^n$ is called an *algebraic set* if there are polynomials $f_1, \ldots, f_r \in \mathbb{K}[T_1, \ldots, T_n]$ such that

$$X = V(f_1, \ldots, f_r) = \{x \in \mathbb{K}^n \mid f_i(x) = 0 \text{ for all } i = 1, \ldots, r\}.$$

The *vanishing ideal* of a subset $X \subseteq \mathbb{K}^n$ is the set of polynomials which vanish on every point of X:

$$I(X) = \{f \in \mathbb{K}[T_1, \ldots, T_n] \mid f(x) = 0 \text{ for each } x \in X\}.$$

The algebraic sets of a $\mathbb{K}^n$ are the closed sets of a topology, called *Zariski topology*. In the following, the concepts of open and closed set are always intended with respect to the Zariski topology, if not explicitly stated otherwise. Note that if $\mathbb{K} = \mathbb{C}$, the Zariski topology is weaker as the usual topology.

The *coordinate ring* of an algebraic set $X \subseteq \mathbb{K}^n$ is

$$\mathbb{K}[X] = \{f : X \to \mathbb{K} \mid f = F_{|X} \text{ for a polynomial } F \in \mathbb{K}[T_1, \ldots, T_n]\}.$$

This is an affine $\mathbb{K}$-algebra, that is, a finitely generated $\mathbb{K}$-algebra with no nilpotent elements.

If $X \subseteq \mathbb{K}^n$ is an algebraic set, the mapping

$$\begin{array}{rccc} \pi_X : & \mathbb{K}[T_1, \ldots, T_n] & \to & \mathbb{K}[X] \\ & F & \mapsto & F_{|X} \end{array}$$

is a surjective morphism and its kernel is $I(X)$. It follows that

$$\mathbb{K}[X] \cong \mathbb{K}[T_1, \ldots, T_n]/I(X),$$

thus $\mathbb{K}[X]$ is a finitely generated $\mathbb{K}$-algebra.

The *principal set* of a function $f \in \mathbb{K}[X]$ is $\{x \in X \mid f(x) \neq 0\} = X \setminus V(f)$.

Let X be an algebraic set. For an ideal $\mathfrak{a} \subseteq \mathbb{K}[X]$ define the *vanishing set in* X

$$V_X(\mathfrak{a}) = \{x \in X \mid f(x) = 0 \text{ for all } f \in \mathfrak{a}\}.$$

For a subset $Y \subseteq X$ define the *vanishing ideal in* $\mathbb{K}[X]$

$$I_X(Y) = \{f \in \mathbb{K}[X] \mid f(y) = 0 \text{ for all } y \in Y\}.$$

A non empty topological space X is said to be irreducible if it can not be decomposed into two closed sets $X_1, X_2 \neq X$. Other characterizations of irreducibility are:

- every two non empty open subset of X have non empty intersection
- every non empty open set is dense in X.

Theorem 2.1.1
A non empty algebraic set $X \subseteq \mathbb{K}^n$ is irreducible if and only if its vanishing ideal $I(X) \subseteq \mathbb{K}[T_1, \ldots, T_n]$ is prime.

The proof can be found in [Milne05], corollary 1.15.

Theorem 2.1.2
A non empty algebraic set $X \subseteq \mathbb{K}^n$ can be decomposed in an unique way as $X = X_1 \cup \ldots \cup X_n$, where each $X_i \subseteq X$ is closed and irreducible, and $X_i \not\subseteq X_j$ if $i \neq j$. The sets X_i are called the irreducible components of X.

For the proof, see [Shaf77], chapter 1, § 3, Theorem 2.

Let $X \subseteq \mathbb{K}^n$ be an algebraic set and $U \subseteq X$ an open set. A function $f : U \to \mathbb{K}$ is said to be *regular* if for each $x \in U$ there are $g, h \in \mathbb{K}[X]$ such that $f = \frac{g}{h}$ in a neighbourhood of x.
The set of the regular functions on U forms the $\mathbb{K}$-*algebra of regular functions on* U. It is denoted by $\mathcal{O}_X(U)$.

A sheaf is a mapping $\mathcal{F}$ from the open sets of a topological space X to the abelian groups (or rings, or $\mathbb{K}$-algebras, ...)

$$\mathcal{F} : \{\text{open subsets of } X\} \to \{\text{abelian groups}\}, \qquad U \mapsto \mathcal{F}(U)$$

such that

(i) for every open sets $V \subseteq W \subseteq X$ there is a restriction homomorphism

$$\rho_V^W : \mathcal{F}(W) \to \mathcal{F}(V), \qquad f \mapsto f_{|V}$$

and if $U \subseteq V \subseteq W$ are open subsets of X, these homomorphisms are compatible,

$$\rho_U^W = \rho_U^V \circ \rho_V^W$$

(ii) if $U = \cup_{i\in I} U_i$, where each $U_i \subseteq X$ is open, then

(a) given $f \in \mathcal{F}(U)$ such that $f_{|U_i} = 0 \in \mathcal{F}(U_i)$, for each $i \in I$, then $f = 0$

(b) given $f_i \in \mathcal{F}(U_i)$ for each i, such that $f_{i|U_i\cap U_j} = f_{j|U_i\cap U_j}$ for each i, j, there is a $f \in \mathcal{F}(U)$ such that $f_{|U_i} = f_i$ for each i.

A mapping which satisfies only the first statement (i) is called a presheaf. For any algebraic set $X \subseteq \mathbb{K}^n$, the mapping $U \mapsto \mathcal{O}_X(U)$, which maps the open subset U of X into the $\mathbb{K}$-algebra of regular functions on U, is a sheaf.

Given a commutative ring R with unity and a multiplicative system $S \subseteq R$ (that means $1 \in S$ and if $a, b \in S$ then also $ab \in S$), we can consider the equivalence relation on $R \times S$

$$(r, s) \sim (r', s') \Leftrightarrow u(rs' - r's) = 0 \text{ for a } u \in S$$

and write $\frac{r}{s}$ for the equivalence class of (r, s). The quotient set is denoted by $S^{-1}R$ and is called fraction ring of R. In particular, for $f \in R$, we define the localization of R by f as $R_f = S^{-1}R$, where $S = \{f^r \mid r \in \mathbb{N}\}$.

Theorem 2.1.3
Let $X \in \mathbb{K}^n$ be an algebraic set and $g \in \mathbb{K}[X]$. Then, the mapping

$$\begin{array}{ccc} \mathbb{K}[X]_g & \to & \mathcal{O}_X(X_g) \\ \frac{f}{g^r} & \mapsto & \left[x \mapsto \frac{f(x)}{g(x)^r}\right] \end{array}$$

is an isomorphism. In particular, taking $g = 1$, one has

$$\mathcal{O}_X(X) = \mathbb{K}[X],$$

i.e. the $\mathbb{K}$-algebra of regular functions is isomorphic to the coordinate ring.

Let X be a topological space and $\mathcal{F}$ a presheaf of abelian groups on it. Consider the disjoint union

$$F = \bigsqcup_{x\in U\subseteq X} \mathcal{F}(U)$$

where U runs through all the open neighbourhoods of x. On F define an equivalence relation as follows: for $s \in \mathcal{F}(U)$ and $s' \in \mathcal{F}(U')$

$s \sim s' \Leftrightarrow$ there is a open neighbourhood $V \subseteq U \cap U'$ such that $s_{|V} = s'_{|V}$

The stalk of $\mathcal{F}$ in a point $x \in X$ is the quotient set

$$\mathcal{F}_x = F/\sim$$

For an element $s \in \mathcal{F}(U)$ and a $x \in U$, we denote by s_x its residue class in $\mathcal{F}_x$ and call it the germ of s in x.

The stalk $\mathcal{F}_x$ has the structure of abelian group: for $s_x, s'_x \in \mathcal{F}_x$ define

$$s_x + s'_x = (s_{|U\cap U'} + s'_{|U\cap U'})_x.$$

In the same way, the stalks of a presheaf of rings (algebras, ...) are again rings (algebras, ...).

Let $\mathfrak{p}$ be a prime ideal of a commutative ring R with unity. Then $S = R\setminus\mathfrak{p}$ is a multiplicative system and we define the localization of R at $\mathfrak{p}$ as

$$R_\mathfrak{p} = S^{-1}R.$$

Theorem 2.1.4
$R_\mathfrak{p}$ is a local ring, that is in $R_\mathfrak{p}$ there is exactly one maximal ideal, which is

$$S^{-1}\mathfrak{p} = \left\{ \frac{r}{s} \mid r \in \mathfrak{p}, s \in S \right\}.$$

Theorem 2.1.5
Let $X \subseteq \mathbb{K}^n$ be an algebraic set and $x \in X$. Denote by $\mathfrak{m}_x = \{f \in \mathbb{K}[X] \mid f(x) = 0\}$ the maximal ideal corresponding to x. Then, the mapping

$$\begin{array}{ccc} \mathbb{K}[X]_{\mathfrak{m}_x} & \to & \mathcal{O}_{X,x} \\ \frac{f}{g} & \mapsto & \left(\frac{f}{g}\right)_x \end{array}$$

from the localization of the coordinate ring onto the stalk $(\mathcal{O}_X(X))_x$, is an isomorphism.

For the proof, see [Milne05], Proposition 3.6.

A *space with functions* is a topological space X together with a sheaf $\mathcal{O}_X$ of $\mathbb{K}$-valued functions.

A continuous function $\varphi : X \to Y$ of spaces with functions is called a *morphism* if for each open set $V \subseteq Y$, for each $f \in \mathcal{O}_Y(V)$, the composition $f \circ \varphi$ lies in $\mathcal{O}_X(\varphi^{-1}(V))$.

Let $\varphi : X \to Y$ be a morphism of spaces with functions and $V \subseteq Y$ an open set. We define the *comorphism*

$$\begin{array}{rcl} \varphi^* : \mathcal{O}_Y(V) & \to & \mathcal{O}_X(\varphi^{-1}(V)) \\ f & \mapsto & f \circ \varphi. \end{array}$$

Theorem 2.1.6
Let $X \in \mathbb{K}^n$ and $Y \in \mathbb{K}^m$ be algebraic sets and $\phi : X \to Y$ be a function. Then the following statements are equivalent:

1. *ϕ is a morphism of spaces with functions*
2. *ϕ is the restriction of a polynomial map $\Phi : \mathbb{K}^n \to \mathbb{K}^m$.*

Definition 2.1.7
An *affine $\mathbb{K}$-variety* is a space with functions which is isomorphic to an algebraic set in some $\mathbb{K}^n$.

Definition 2.1.8
Let F be a covariant (contravariant) functor of the category $\mathcal{C}$ in the category $\mathcal{D}$. F is an equivalence if the following two conditions hold:

- for each object $Y \in \mathcal{D}$ there is an object $X \in \mathcal{C}$ such that $F(X) = Y$;
- for every two objects $X, X' \in \mathcal{C}$ the mapping

$$\begin{array}{rcl} \mathrm{Mor}(X, X') & \to & \mathrm{Mor}(F(X), F(X')) \\ \phi & \mapsto & F(\phi) \end{array}$$

  is a bijection. (The codomain is $\mathrm{Mor}(F(X'), F(X))$ in case of a contravariant functor.)

Theorem 2.1.9 ("Anti-equivalence" Theorem)
Let X be an algebraic set. The functor which maps

$$X \mapsto \mathbb{K}[X] \qquad \text{and} \qquad \varphi \mapsto \varphi^*$$

is a contravariant equivalence from the category of the algebraic sets to the category of the affine $\mathbb{K}$-algebras. That is,

- *it is a contravariant functor;*
- *for each affine algebra A there is an $n \in \mathbb{N}$ and an algebraic set $X \subseteq \mathbb{K}^n$, such that A is isomorphic to $\mathbb{K}[X]$;*
- *if $X \subseteq \mathbb{K}^n$ and $Y \subseteq \mathbb{K}^m$ are algebraic sets, then there is a bijection*

$$\{\text{morphisms } X \to Y\} \to \{\text{homomorphisms } \mathbb{K}[Y] \to \mathbb{K}[X]\}$$
$$\varphi \mapsto \varphi^*.$$

Theorem 2.1.10
Let X be an affine variety and $f \in \mathcal{O}(X)$. Then the principal set $X_f = \{x \in X \mid f(x) \neq 0\}$ with the subspace topology and the induced sheaf is an affine variety, and $\mathcal{O}(X_f) = \mathcal{O}(X)_f$.

2.2 Affine tangent spaces

Let $f \in \mathbb{K}[T_1, \ldots, T_n]$ be a polynomial and $a \in \mathbb{K}^n$ a point. Define

$$L_{f,a} \;=\; \sum_{i=1}^{n} \frac{\partial f}{\partial T_i}(a)\,(T_i - a_i) \;=\; \langle \mathrm{grad}_f(a), T - a \rangle$$

The *affine tangent space to an algebraic set X in a point $a \in X$* is the space

$$T_a^{\mathrm{aff}} X \;=\; V(L_{f,a} \mid f \in I(X))$$

We give $T_a^{\mathrm{aff}} X$ the structure of vector space: for $z = a + v$, $z' = a + v'$, and a scalar $c \in \mathbb{K}$, define

$$z + z' = a + v + v', \qquad cz = a + cv,$$

that means also, a is the zero of $T_a^{\mathrm{aff}} X$.

Proposition 2.2.1
If $f_1, \ldots, f_r$ generate $I(X)$, then $T_a^{\mathrm{aff}} X \;=\; V(L_{f_i,a} \mid i = 1, \ldots, r)$ and the dimension of the affine tangent space is linked with the rank of the Jacobi matrix by the relation

$$\dim(T_a^{\mathrm{aff}} X) = n - \mathrm{rk} \left(\frac{\partial f_i}{\partial T_j}(a) \right)_{\substack{1 \le i \le r \\ 1 \le j \le n}}$$

Proof. We have to prove that $\{L_{f_1,a}, \ldots, L_{f_r,a}\}$ generate the same ideal in the polynomial ring $\mathbb{K}[T_1, \ldots, T_n]$ as the set $\{L_{f,a} \mid f \in I(X)\}$, that means, we have to write each $L_{f,a}$ as linear combination of the $L_{f_i,a}$. To do that, note that for each $f \in I(X)$ there are $h_i \in \mathbb{K}[T_1, \ldots, T_n]$ such that $f = \sum h_i f_i$, because $f_1, \ldots, f_r$ generate $I(X)$. Then

$$\begin{aligned} L_{f,a} &= \langle \mathrm{grad}_f(a), T - a \rangle \\ &= \sum \langle \mathrm{grad}_{h_i}(a) f_i(a), T - a \rangle + \sum \langle h_i(a) \mathrm{grad}_{f_i}(a), T - a \rangle \\ &= \sum \langle h_i(a) \mathrm{grad}_{f_i}(a), T - a \rangle \\ &= \sum h_i(a) L_{f_i,a}. \end{aligned}$$

To prove the connection of the dimension with the Jacobi matrix, for each

point $a \in \mathbb{K}^n$ consider the matrix

$$\begin{aligned} M_a &= (L_{f_i,a}(e_j + a))_{\substack{1 \le i \le r \\ 1 \le j \le n}} \\ &= \left(\sum_{k=1}^{n} \frac{\partial f_i}{\partial T_j}(a)(T_k(e_j + a) - a_k) \right)_{i,j} \\ &= \left(\frac{\partial f_i}{\partial T_j}(a) \right)_{\substack{1 \le i \le r \\ 1 \le j \le n}} \end{aligned}$$

If $a \in X$, then the dimension of $T_a^{\mathrm{aff}} X$ is the dimension of the kernel of M_a, and the proof is complete. □

Remark 2.2.2
Let $X \subseteq \mathbb{K}^n$ be an algebraic set. For each $n \in \mathbb{N}$, the set

$$A_k = \{x \in X \mid \dim(T_x^{\mathrm{aff}} X) \ge n\}$$

is closed in X.
In particular, the points whose tangent space has minimal dimension form a nonempty open set in X.

Proof. Since

$$\begin{aligned} A_k &= \{x \in X \mid \mathrm{rk}(M_x) \le n - k\} \\ &= \{x \in X \mid \det(N_x) = 0 \text{ for all } (n-k+1)\text{-minors } N_x \text{ of } M_x\}, \end{aligned}$$

A_k is the intersection of X with an algebraic set in $\mathbb{K}^n$, hence it is closed. The second assertion follows immediately. □

Definition 2.2.3
The Zariski dimension of an irreducible algebraic set X is defined through the dimension of the affine tangent space,

$$\mathrm{Zdim}\, X = \min_{x \in X} \dim T_x^{\mathrm{aff}} X.$$

The Zariski dimension of an algebraic set is the maximum Zariski dimension of its irreducible components.

In the next section, we extend the notion of Zariski dimension to irreducible affine varieties, see Definition 2.3.7.

2.3 Derivations

Let R be a $\mathbb{K}$-algebra with unity, and $\alpha : R \to \mathbb{K}$ a homomorphism. An *α-derivation* is a $\mathbb{K}$-linear map $\delta : R \to \mathbb{K}$ such that for each $f, g \in R$ the following product derivation rule holds:

$$\delta(fg) = \delta(f)\alpha(g) + \alpha(f)\delta(g).$$

The set of α-derivations is a vector space on $\mathbb{K}$; we denote it by $\mathrm{Der}_\alpha(R)$.

With abuse of notation, we identify $1 \cdot \mathbb{K} \subseteq R$ with the field $\mathbb{K}$ and we write $\alpha(\alpha(g)) = \alpha(g)$ and $\alpha(1) = 1$.

Let $X \subseteq \mathbb{K}^n$ be an algebraic set and $x \in X$ a point. Let α be the evaluation homomorphism, $\alpha : \mathbb{K}[X] \to \mathbb{K}$, $f \mapsto f(x)$. Its kernel is the maximal ideal $\mathfrak{m}_x \subseteq \mathbb{K}[X]$. The vector space of the α-derivations of $\mathbb{K}[X]$ is $\mathrm{Der}_x(\mathbb{K}[X])$ and its elements are called the derivations of $\mathbb{K}[X]$ centered at x.

Let $z = x + v \in T_x^{\mathrm{aff}} X$. Assign to z the following derivation of the polynomial ring $\mathbb{K}[T_1, \ldots, T_n]$ centered at x:

$$z \;\mapsto\; \left[\rho_z : f \;\mapsto\; L_{f,x}(z) = \langle \mathrm{grad}_f(x), v \rangle\right].$$

The usual properties of derivatives guarantee that ρ_z is actually a derivation and since $z \in T_x^{\mathrm{aff}} X$, this derivation vanishes on each element of $I(X)$, and so it induces a derivation $\delta_z \in \mathrm{Der}_x(\mathbb{K}[X])$ with $\delta_z(f_{|X}) = \rho_z(f)$. δ_z is called the directional derivative associated to z.
Define Φ the canonical epimorphism

$$\Phi : \mathbb{K}[T_1, \ldots, T_n] \to \mathbb{K}[X], \qquad f \mapsto f_{|X},$$

then we have $\rho_z = \delta_z \circ \Phi$.

Proposition 2.3.1
Let R be a $\mathbb{K}$-algebra, $\alpha : R \to \mathbb{K}$ a homomorphism, and $\delta \in \mathrm{Der}_\alpha(R)$. Suppose $\mathfrak{m} = \ker \alpha$. Then

- *for each $k \in \mathbb{K}$, we have $\delta(k) = 0$;*
- *for each $f \in \mathfrak{m}^2$, $\delta(f) = 0$.*

Proof. We have $\delta(1) = 0$, since

$$\delta(1) = \delta(1 \cdot 1) = \delta(1)\alpha(1) + \alpha(1)\delta(1) = 2\delta(1).$$

Hence, and since δ is $\mathbb{K}$-linear, follows that $\delta(k \cdot 1) = k\delta(1) = 0$.
For the second part, suppose $f \in \mathfrak{m}^2$; then it can be written as $\sum h_i g_i$ with

suitable $g_i, h_i \in \mathfrak{m}$. Thus, by linearity and the product derivation rule, we have

$$\delta(f) = \delta\left(\sum h_i g_i\right) = \sum \left(\delta(g_i)\alpha(h_i) + \delta(h_i)\alpha(g_i)\right) = 0,$$

since $\alpha(h_i) = \alpha(g_i) = 0$ for all i. □

Theorem 2.3.2
Let $X \subseteq \mathbb{K}^n$ be an algebraic set and $a \in X$. Then the mapping

$$T_a^{\text{aff}} X \to \text{Der}_a(\mathbb{K}[X]) \qquad z \mapsto \delta_z$$

is an isomorphism between vector spaces.

Proof. The mapping $z \mapsto \delta_z$ is a homomorphism, since for $z = a + v, z' = a + v' \in T_a^{\text{aff}} X$, we have

- $\rho_{z+z'} : f \to L_{f,a}(z + z') = \sum \frac{\partial f}{\partial T_i}(a)(v_i + v_i') = L_{f,a}(z) + L_{f,a}(z')$.
- $\rho_{cz} : f \to L_{f,a}(cz) = \sum \frac{\partial f}{\partial T_i}(a)(cv_i) = cL_{f,a}(z)$.
- $\rho_e : f \to L_{f,a}(e) = \sum \frac{\partial f}{\partial T_i}(a)$, where e denotes the unity in $T_a^{\text{aff}} X$.

Let $\Phi : \mathbb{K}[T_1, \ldots, T_n] \to \mathbb{K}[X]$ be the canonical epimorphism defined by $f \mapsto f_{|X}$.

We first show that $z \mapsto \delta_z$ is surjective: for that purpose let $\delta : \mathbb{K}[X] \to \mathbb{K}$ be a derivation centered at a, and look for some z such that $\delta = \delta_z$. Let $\delta' = \delta \circ \Phi$; it is a derivation centered at a of the polynomial ring $\mathbb{K}[T_1, \ldots, T_n]$. Define $v = (\delta'(T_1), \ldots, \delta'(T_n))$ and $z = v + a$. We claim that $z \in T_a^{\text{aff}} X$. To prove this, we must show that for each $g \in I(X)$, the linear part $L_{g,a}$ vanishes in z. We have

$$\delta'(g) = \delta(\Phi(g)) = \delta(0) = 0.$$

The Taylor expansion of g in a yields $g = g(a) + L_{g,a} + h$, with $h \in \mathfrak{m}^2$, where $\mathfrak{m}$ is the maximal ideal $\{f \in \mathbb{K}[X] \mid f(a) = 0\}$. By proposition 2.3.1, applying δ' to the Taylor expansion, we have

$$0 = \delta'(g) = \delta'(L_{g,a}) = \sum_{i=1}^{n} \frac{\partial g}{\partial T_i}(a)\delta'(T_i) = \sum_{i=1}^{n} \frac{\partial g}{\partial T_i}(a)v_i = L_{g,a}(z).$$

Thus, $z \in T_a^{\text{aff}} X$ is verified. The derivations $\delta' = \delta \circ \Phi$ and $\delta_z \circ \Phi$ coincide by construction. Since Φ is surjective, it follows that $\delta = \delta_z$, hence $z \mapsto \delta_z$ is surjective.

To prove that $z \mapsto \delta_z$ is injective, let $z = v+a \in T_a^{\text{aff}}X$, with $\delta_z = 0$. Then $\rho_z = \delta_z \circ \Phi$ is a derivation centered at a of the polynomial ring $\mathbb{K}[T_1, \ldots, T_n]$ and for each i

$$0 = \rho_z(T_i) = \sum_{j=1}^{n} \frac{\partial T_i}{\partial T_j}(a)v_j = \frac{\partial T_i}{\partial T_i}(a)v_i = v_i.$$

Hence, $v = 0$. Due to the definition of the structure of vector space on $T_a^{\text{aff}}X$, we have that $z = a$ is the zero of $T_a^{\text{aff}}X$. □

Let X be an affine variety, $x \in X$ and $\alpha : \mathcal{O}_{X,x} \to \mathbb{K}$ the evaluation homomorphism

$$\begin{aligned} \alpha : \mathcal{O}_{X,x} &\to \mathbb{K} \\ f_x &\mapsto f(x). \end{aligned}$$

It is well defined, because $f \in \mathcal{O}(U)$ and $f' \in \mathcal{O}(U')$ are equivalent if and only if $f_{|V} = f'_{|V}$ for some open set $V \subseteq U \cap U'$ such that $x \in V$.

The $\mathbb{K}$-vector space of the derivations on $\mathcal{O}_{X,x}$ is defined as

$$\text{Der}(\mathcal{O}_{X,x}) = \text{Der}_\alpha(\mathcal{O}_{X,x}).$$

Thus, a derivation on the stalk $\mathcal{O}_{X,x}$ is a $\mathbb{K}$-linear mapping $\delta : \mathcal{O}_{X,x} \to \mathbb{K}$ such that the product rule holds: $\delta(fg) = \delta(f)g(x) + f(x)\delta(g)$.

We extend the concept of affine tangent space, which is so far defined only for algebraic sets, also to affine variety.

Definition 2.3.3
The *affine tangent space to an affine variety X in a point $x \in X$* is defined as

$$T_{X,x} = \text{Der}(\mathcal{O}_{X,x}).$$

We will show in 2.3.6 that for every algebraic set X the affine tangent space $T_x^{\text{aff}}X$ is isomorphic to the affine tangent space $T_{X,x}$ of X seen as an affine variety.

Proposition 2.3.4
Let R be a local $\mathbb{K}$-algebra with maximal ideal $\mathfrak{m} \subseteq R$. Then there is a decomposition $R = \mathbb{K} \oplus \mathfrak{m}$ and the mapping

$$\begin{aligned} \alpha_{\mathfrak{m}} : \quad R = \mathbb{K} \oplus \mathfrak{m} &\to \mathbb{K} \\ f = a \oplus f' &\mapsto a \end{aligned}$$

is the unique homomorphism of algebras of R in $\mathbb{K}$.

As in the case of the stalk of the sheaf $\mathcal{O}_X$, we define the vector space of derivations on R:

$$\mathrm{Der}(R) = \mathrm{Der}_{\alpha_{\mathfrak{m}}}(R).$$

Theorem 2.3.5
Let R be a $\mathbb{K}$-algebra, $\alpha : R \to \mathbb{K}$ a homomorphism of $\mathbb{K}$-algebras and $\mathfrak{m} = \ker(\alpha)$. Then there is a canonical isomorphism

$$\begin{array}{rccl} \varphi: & \mathrm{Der}_\alpha(R) & \to & \mathrm{Der}(R_{\mathfrak{m}}) \\ & \delta & \mapsto & \left[\frac{r}{s} \mapsto \frac{\delta(r)\alpha(s)-\alpha(r)\delta(s)}{\alpha(s)^2}\right]. \end{array}$$

Proof. We need show that

(1) φ is well defined;

(2) $\varphi(\delta)$ is a derivation of $R_{\mathfrak{m}}$;

(3) φ is an isomoprhism.

To prove (1), let $r, r' \in R$ and $s, s' \in R \setminus \mathfrak{m}$ such that $\frac{r}{s} = \frac{r'}{s'}$. By definition of $R_{\mathfrak{m}}$ there is a $u \in R \setminus \mathfrak{m}$ such that $u(rs' - r's) = 0$. Then

$$0 = \alpha(0) = \alpha(u)\alpha(rs' - r's)$$

and since u is not in the kernel of α, this holds if and only if

$$0 = \alpha(rs' - r's) = \alpha(r)\alpha(s') - \alpha(r')\alpha(s). \tag{2.3.1}$$

Furthermore, applying δ and the product rule to the equality $u(rs'-r's) = 0$, we get

$$\begin{array}{rcl} 0 & = & \delta(u(rs' - r's)) \\ & = & \delta(u)\alpha(rs' - r's) + \alpha(u)\delta(rs' - r's) \\ & = & \delta(u)(\alpha(r)\alpha(s') - \alpha(r')\alpha(s))+ \\ & & +\alpha(u)\left(\delta(r)\alpha(s') + \alpha(r)\delta(s') - \delta(r')\alpha(s) - \alpha(r')\delta(s)\right), \end{array}$$

where the first summand is 0 by 2.3.1. Hence, since $\alpha(u) \neq 0$, we obtain

$$\delta(r)\alpha(s') - \alpha(r')\delta(s) = \delta(r')\alpha(s) - \alpha(r)\delta(s'). \tag{2.3.2}$$

By 2.3.1, we have

$$\alpha(r') = \frac{\alpha(r)\alpha(s')}{\alpha(s)} \qquad \text{and} \qquad \alpha(r) = \frac{\alpha(r')\alpha(s)}{\alpha(s')},$$

hence, substituting these expressions in 2.3.2,

$$\alpha(s')\left(\delta(r)-\frac{\alpha(r)}{\alpha(s)}\delta(s)\right)=\alpha(s)\left(\delta(r')-\frac{\alpha(r')}{\alpha(s')}\delta(s')\right),$$

which gives

$$\frac{\delta(r)\alpha(s)-\alpha(r)\delta(s)}{\alpha(s)^2}=\frac{\delta(r')\alpha(s')-\alpha(r')\delta(s')}{\alpha(s')^2}.$$

That means $\varphi(\delta)\left(\frac{r}{s}\right)=\varphi(\delta)\left(\frac{r'}{s'}\right)$ if $\frac{r}{s}=\frac{r'}{s'}$, hence φ does not depend on the choice of the class representative, that is it is well defined.

Now we show (2). Clearly $\varphi(\delta)$ is $\mathbb{K}$-linear and vanishes on $\mathbb{K}$. To prove that the product rule holds, calculate

$$\begin{aligned}
\varphi(\delta)\left(\frac{r}{s}\frac{r'}{s'}\right)&=\frac{\delta(rr')\alpha(ss')-\delta(ss')\alpha(rr')}{\alpha(ss')^2}\\
&=\frac{\left(\delta(r)\alpha(r')+\delta(r')\alpha(r)\right)\alpha(ss')\;-\;\left(\delta(s)\alpha(s')+\delta(s')\alpha(s)\right)\alpha(rr')}{\alpha(ss')^2}\\
&=\frac{\delta(r)\alpha(s)-\delta(s)\alpha(r)}{\alpha(s)^2}\,\frac{\alpha(r')}{\alpha(s')}\;+\;\frac{\delta(r')\alpha(s')-\delta(s')\alpha(r')}{\alpha(s')^2}\,\frac{\alpha(r)}{\alpha(s)}\\
&=\frac{\alpha(r')}{\alpha(s')}\,\varphi(\delta)\left(\frac{r}{s}\right)\;+\;\frac{\alpha(r)}{\alpha(s)}\,\varphi(\delta)\left(\frac{r'}{s'}\right)\\
&=\alpha\left(\frac{r'}{s'}\right)\varphi(\delta)\left(\frac{r}{s}\right)\;+\;\alpha\left(\frac{r}{s}\right)\varphi(\delta)\left(\frac{r'}{s'}\right).
\end{aligned}$$

Hence, $\varphi(\delta)$ is a derivation of $R_{\mathfrak{m}}$.

To prove (3), note that φ is a homomorphism of $\mathbb{K}$-vector fields.
To show that φ is injective, suppose $\varphi(\delta)=0$. Then for each $r\in R$

$$0=\varphi(\delta)\left(\frac{r}{1}\right)=\delta(r),$$

hence $\delta=0$, and φ is injective.
To show that φ is surjective, take a $\delta'\in\mathrm{Der}(R_{\mathfrak{m}})$. Then, $\delta(r)=\delta'\left(\frac{r}{1}\right)$ defines an α-derivation on R and

$$\begin{aligned}
\varphi(\delta)\left(\frac{r}{s}\right)&=\frac{\delta'(\frac{r}{1})\alpha(s)-\alpha(r)\delta'(\frac{s}{1})}{\alpha(s)^2}\\
&=\delta'\left(\frac{r}{1}\right)\alpha\left(\frac{1}{s}\right)-\alpha\left(\frac{r}{1}\right)\delta'\left(\frac{s}{1}\right)\alpha\left(\frac{1}{s^2}\right)
\end{aligned}$$

Since

$$0 = \delta'\left(\frac{s}{s}\right) = \delta'\left(\frac{s}{1}\right)\alpha\left(\frac{1}{s}\right) + \delta'\left(\frac{1}{s}\right)\alpha\left(\frac{s}{1}\right),$$

we have $\delta'\left(\frac{1}{s}\right) = -\delta'\left(\frac{s}{1}\right)\alpha\left(\frac{1}{s^2}\right)$. Hence

$$\varphi(\delta)\left(\frac{r}{s}\right) = \delta'\left(\frac{r}{1}\right)\alpha\left(\frac{1}{s}\right) + \alpha\left(\frac{r}{1}\right)\delta'\left(\frac{1}{s}\right) = \delta'\left(\frac{r}{s}\right),$$

which concludes the proof that φ is an isomorphism. □

Proposition 2.3.6
Let X be an algebraic set in $\mathbb{K}^n$ and $x \in X$ a point.Then there are canonical isomorphisms

$$T_x^{\text{aff}}X \;\cong\; \text{Der}_x(\mathcal{O}(X)) \;\cong\; \text{Der}(\mathcal{O}_{X,x}) \;=\; T_{X,x}.$$

Proof. By Theorem 2.3.2, $T_x^{\text{aff}}X \cong \text{Der}_x(\mathbb{K}[X])$ and by Theorem 2.1.3, $\mathcal{O}_X(X) = \mathbb{K}[X]$. By Theorem 2.3.5, $\text{Der}_x(\mathcal{O}(X)) \cong \text{Der}(\mathcal{O}_X(X)_{\mathfrak{m}_x})$, and since $\mathcal{O}_X(X)_{\mathfrak{m}_x}$ is isomorphic to $\mathcal{O}_{X,x}$, the proof is complete. □

Thanks to the previous result, we can now extend the definition of Zariski dimension given in Definition 2.2.3 to affine varieties.

Definition 2.3.7
Let X be an irreducible affine variety. Its Zariski dimension is defined as

$$\text{Zdim}\, X = \min_{x \in X} \dim T_{X,x}.$$

The Zariski dimension of an affine variety is the maximum Zariski dimension of its irreducible components.

Definition 2.3.8
Let X be an affine variety and $x \in X$ be a point. Then the dual space $(\mathfrak{m}_x/\mathfrak{m}_x^2)^*$ is called Zariski tangent space of X at x.

Theorem 2.3.9
Let R be a $\mathbb{K}$-algebra, $\alpha : R \to \mathbb{K}$ a homomorphism of $\mathbb{K}$-algebras and $\mathfrak{m} = \ker(\alpha)$. Then the mapping

$$\begin{array}{rcl} \text{Der}_\alpha(R) & \to & (\mathfrak{m}/\mathfrak{m}^2)^* \\ \delta & \mapsto & [u_\delta : f + \mathfrak{m}^2 \mapsto \delta(f)] \end{array}$$

is an isomorphism of $\mathbb{K}$-vector spaces.

Proof. By Proposition 2.3.1, the map $\delta \mapsto u_\delta$ is well defined. It is clearly a linear map.

To show that the mapping is injective, let $\delta \in \mathrm{Der}_\alpha(R)$ such that $u_\delta = 0$, that means $0 = u_\delta(f + \mathfrak{m}^2) = \delta(f)$ for each $f \in \mathfrak{m}$. We have to show that $\delta = 0$, i.e. $\delta(f) = 0$ for all $f \in R$. We have, by the linearity of δ,

$$\delta(f) = \delta(f - \alpha(f) + \alpha(f)) = \delta(f - \alpha(f)) + \delta(\alpha(f)).$$

Now, $\delta(\alpha(f)) = 0$ because $\alpha(f) \in \mathbb{K}$ and since α is $\mathbb{K}$-linear, $\alpha(f - \alpha(f)) = \alpha(f) - \alpha(\alpha(f)) = 0$, that means $f - \alpha(f) \in \ker \alpha = \mathfrak{m}$, thus $\delta(f - \alpha(f)) = 0$. Hence, $\delta(f) = 0$ for all $f \in R$.

To show that the mapping is surjective, let $u : \mathfrak{m}/\mathfrak{m}^2 \to \mathbb{K}$ be a linear functional. Define the $\mathbb{K}$-linear map

$$\begin{array}{rccl} \delta: & R & \to & \mathbb{K} \\ & f & \mapsto & u(f - \alpha(f) + \mathfrak{m}^2). \end{array}$$

Let us prove that it is an α-derivation: first note that δ vanishes on $\mathbb{K}$ and on $\mathfrak{m}^2$.
Furthermore, for each $h \in R$, we have

$$\begin{aligned} \delta(h - \alpha(h)) &= u(h - \alpha(h) - \alpha(h - \alpha(h)) + \mathfrak{m}^2) \\ &= u(h - \alpha(h) + \mathfrak{m}^2) \\ &= \delta(h). \end{aligned}$$

Now, to obtain the product rule, consider

$$\begin{aligned} \delta(fg) &= \delta(fg - \alpha(fg)) \\ &= \delta\big((f - \alpha(f))\alpha(g) + \alpha(f)(g - \alpha(g)) + (f - \alpha(f))(g - \alpha(g))\big) \\ &= \delta((f - \alpha(f))\alpha(g)) + \delta((g - \alpha(g))\alpha(f)) + \delta((f - \alpha(f))(g - \alpha(g))) \\ &= \delta(f)\alpha(g) + \delta(g)\alpha(f), \end{aligned}$$

since $(f - \alpha(f))(g - \alpha(g)) \in \mathfrak{m}^2$.

Now, to prove that $\delta \mapsto u_\delta$ is surjective, note that for each $f + \mathfrak{m}^2 \in \mathfrak{m}/\mathfrak{m}^2$, by construction we have

$$u_\delta(f + \mathfrak{m}^2) = \delta(f) = u(f - \alpha(f) + \mathfrak{m}^2) = u(f + \mathfrak{m}^2) - u(\alpha(f) + \mathfrak{m}^2) = u(f + \mathfrak{m}^2).$$

□

By Theorem 2.1.5, the mapping

$$O_X(X)_{\mathfrak{m}_x} \to \mathcal{O}_{X,x} \quad \frac{f}{g} \mapsto \left(\frac{f}{g}\right)_x$$

is an isomorphism, hence the stalk $\mathcal{O}_{X,x}$ is a local ring, and its maximal ideal is

$$\mathfrak{m}_{X,x} = \left\{ \left(\frac{f}{g}\right)_x \mid f \in \mathfrak{m}_x, g \in \mathcal{O}_X(X) \setminus \mathfrak{m}_x \right\} = \{h_x \in \mathcal{O}_{X,x} \mid h(x) = 0\}.$$

Proposition 2.3.10
Let X be an affine variety and $x \in X$ be a point.Then there are canonical isomorphisms

$$\begin{aligned} T_{X,x} &= \operatorname{Der}(\mathcal{O}_{X,x}) \cong \left(\mathfrak{m}_{X,x}/\mathfrak{m}_{X,x}^2\right)^* \\ &\cong \operatorname{Der}_x(\mathcal{O}_X(X)) \cong \left(\mathfrak{m}_x/\mathfrak{m}_x^2\right)^*. \end{aligned}$$

Proof. Immediate by Proposition 2.3.6 and Theorem 2.3.9. □

2.4 Regular and singular points

We introduce the concept of regular points.

Definition 2.4.1
A point x of an affine variety X is called *smooth* or *regular* if there is an open neighbourhood U of x such that $\dim(T_{X,u}) = \dim(T_{X,x})$ for each $u \in U$.
The non-regular points are called *singular*.

Proposition 2.4.2
Let $X \subseteq \mathbb{K}^n$ be an algebraic set, $I(X) = \langle f_1, \ldots, f_r \rangle$ its vanishing ideal and $x_0 \in X$. Then, the following statements are equivalent:

1. *x_0 is regular*
2. *$x \mapsto \dim(T_{X,x})$ has a local minimum at x_0*
3. *the rank of the Jacobi matrix $\left(\frac{\partial f_i}{\partial T_j}(x)\right)$ has a local maximum at x_0.*

Proof. The equivalence of 2 and 3 is clear by virtue of Proposition 2.2.1 and Remark 2.2.2.
It is clear that 1 implies 2.
To show that 2 implies 1, let U be an open neighbourhood of x_0, such that

$\dim(T_{X,x}) = k$ is a local minimum in U of the mapping $x \mapsto \dim(T_{X,x})$. By Remark 2.2.2 the set $A_k = \{x \in X \mid \dim(T_{X,x}) \geq k+1\}$ is closed in X. Hence, $U \setminus A$ is an open neighbourhood of x_0 in X, and $\dim(T_{X,x})$ is constant on $U \setminus A$. □

Proposition 2.4.3
A point x of an irreducible affine variety X is regular if and only if $\dim(T_{X,x}) = \mathrm{Zdim}(X)$.

Proof. Let $\mathrm{Zdim}(X) = n$. By Remark 2.2.2 the following set is open and nonempty:

$$V = \{p \in X \mid \dim(T_{X,p}) = n\} \subseteq X.$$

Suppose $\dim(T_{X,x}) = n$. Then $x \in V$, which is open, thus x is regular.

Conversely, suppose x is regular, then there is an open set $U \subseteq X$ such that for each $u \in U$ the dimensions $\dim(T_{X,x}) = \dim(T_{X,u})$ coincide. Since X is irreducible, each two open subsets of X have nonempty intersection. Then, $U \cap V \neq \emptyset$. Hence, there is a $u \in U \cap V$, and, since $\dim(T_{X,u}) = n$, we can conclude $\dim(T_{X,x}) = n$. □

Note that this does not hold in general if the affine variety is reducible. As an example consider the variety in $\mathbb{C}^3$ defined by $X = V(z_1z_2, z_1z_3)$. Its irreducible components are $X_1 = V(z_1)$ and $X_2 = V(z_2, z_3)$. All points are regular except for the origin, but the dimension of the tangent space at the regular points of X_1 is 2, while the dimension of the tangent space at the regular points of X_2 is 1.

Theorem 2.4.4
The set of regular points of an affine variety X is open and dense in X.

Proof. By definition, the set X^{reg} of regular points is open. We need only show that it is dense in X.

Let $X_1, \ldots, X_r$ be the irreducible components of X. We claim that each set $U_i = X^{\mathrm{reg}} \cap X_i$ is dense in X_i. Consider

$$V_i = X \setminus \bigcup_{i \neq j} X_j \subseteq X_i$$

It is not empty, because otherwise we could decompose X in $r-1$ irreducible components. Moreover, V_i is open in X.

Let $V_i' \subseteq X_i$ be the set of points $x \in X_i$ for which $\dim(T_{X_i,x})$ is minimal in X_i. V_i' is not empty and open in X_i.

Let $V_i'' = V_i \cap V_i'$. Because of the irreducibility of X_i, since V_i and V_i' are open in X_i, the set V_i'' is not empty; it is also open in X. Then $\mathcal{O}_{X_i}(V) =$

$\mathcal{O}_X(V)$ for each open subset $V \subseteq V_i''$. As a consequence, $\mathcal{O}_{X_i,x} \cong \mathcal{O}_{X,x}$ for all $x \in V_i''$.

Then for each point x of V_i'', $\dim(T_{X_i,x}) = \dim(T_{X,x})$. By definition of V_i'', on this set the function $x \mapsto \dim(T_{X,x})$ is constant. Then $V_i'' \subseteq X^{\text{reg}} \cap X_i = U_i$.

As a consequence, U_i is not empty, and since it is open in X_i, which is irreducible, U_i is dense in X_i.

Finally,

$$X = \bigcup_{i=1}^{r} X_i = \bigcup_{i=1}^{r} \overline{U_i} \subseteq \overline{X^{\text{reg}}}$$

This proves that the set of regular points is dense in X. □

2.5 Transcendence degree

Let $\mathbb{K} \subseteq \mathbb{L}$ be a field extension. An element $l \in \mathbb{L}$ is said to be *algebraic over* $\mathbb{K}$, if there is a non zero polynomial $f \in \mathbb{K}[T]$ such that $f(l) = 0$. If $l \in \mathbb{L}$ is not algebraic, it is said to be *transcendent*.
A field extension $\mathbb{K} \subseteq \mathbb{L}$ is *algebraic* if every element $l \in \mathbb{L}$ is algebraic over $\mathbb{K}$.

A subset $L \in \mathbb{L}$ is said to be *algebraically dependent over* $\mathbb{K}$ if there exist a positive integer n, distinct elements $l_1, \ldots, l_n \in L$ and a non zero polynomial $f \in \mathbb{K}[T_1, \ldots, T_n]$ such that $f(l_1, \ldots, l_n) = 0$. Otherwise L is said to be *algebraically independent over* $\mathbb{K}$.

A *transcendence basis* of $\mathbb{K} \subseteq \mathbb{L}$ is an algebraically independent maximal set $B \subseteq \mathbb{L}$. It is a well known fact that every two transcendence basis of the same field extension have the same cardinality.

Every field extension $\mathbb{K} \subseteq \mathbb{L}$ has a transcendence basis. The *transcendence degree* of the extension is the cardinality of one of its basis B, $\text{trdeg}_{\mathbb{K}}(\mathbb{L}) = |\,B\,|$.

For any irreducible affine variety X, the ideal $I(X)$ is prime, hence the coordinate ring $\mathbb{K}[X] \cong \mathcal{O}_X(X)$ is an integral domain, and we can speak of the quotient field of $\mathcal{O}_X(X)$ and give the following definition.

Definition 2.5.1
The *field of rational functions* $\mathbb{K}(X)$ of an irreducible affine variety X is the quotient field of the ring $\mathcal{O}_X(X)$.

Definition 2.5.2
The *dimension* of an irreducible affine variety X is defined as the transcendence degree of the field of rational functions,

$$\dim(X) = \text{trdeg}_{\mathbb{K}}(\mathbb{K}(X)).$$

The *dimension* of an affine variety X with irreducible components $X_1, \ldots, X_r$ is the maximum of the dimensions of the irreducible components,

$$\dim(X) = \max_{1 \leq i \leq r} \dim(X_i).$$

For example, the affine variety $\mathbb{K}^n$ has dimension n, because its function field is $\mathbb{K}(T_1, \ldots, T_n)$ and a transcendence basis for the extension $\mathbb{K} \subseteq \mathbb{K}(T_1, \ldots, T_n)$ is $\{T_1, \ldots, T_n\}$.

Note that the transcendence degree of $\mathbb{K} \subseteq \mathbb{K}(X)$ is always finite and so is the dimension, because $\mathbb{K}(X) = \mathbb{K}(T_{1|X}, \ldots, T_{n|X})$.

Proposition 2.5.3
An affine variety X is finite if and only if $\dim(X) = 0$.

Proof. Suppose X is finite and let $X_1, \ldots, X_r$ be the irreducible components of X. Then each X_i consists of only one point (because it is irreducible). Hence, we need only show that each affine variety $Y = \{y\}$ consisting of exactly one point is zero dimensional. Since obviously $\mathcal{O}(Y) = \mathbb{K}$, we have $\mathbb{K}(Y) = \mathbb{K}$, and $\operatorname{trdeg}_{\mathbb{K}}(\mathbb{K}(Y)) = 0$.
Conversely, if X has dimension zero, we can assume X is irreducible. Otherwise, we take one irreducible component with maximal dimension; we show that it is zero dimenzional, hence the other components must also be zero dimensional. Since $\operatorname{trdeg}_{\mathbb{K}}(\mathbb{K}(X)) = 0$, the extension $\mathbb{K} \subseteq \mathbb{K}(X)$ is algebraic, hence, because $\mathbb{K}$ is algebraically closed, $\mathbb{K}(X) = \mathbb{K}$. It follows that $\mathbb{K}[X] = \mathbb{K}$.
By the Anti-equivalence Theorem 2.1.9, each morphism $Y \to X$ corresponds bijectively to one homomorphism $\mathbb{K}[X] \to \mathbb{K}[Y]$. Now, taking a variety $Y = \{y\}$ containing exactly one point, we have $\mathbb{K}[Y] = \mathbb{K} = \mathbb{K}[X]$, hence there is only one homomorphism $\mathbb{K}[X] \to \mathbb{K}[Y]$. Then the Anti-equivalence Theorem ensures that there is only one morphism $Y \to X$, that means X consists of exactly one point. □

Theorem 2.5.4
Let X be an affine variety and $Y \subseteq X$ a closed subvariety. Then

$$\dim(Y) \leq \dim(X).$$

Proof. Without loss of generality we may suppose X and Y to be irreducible. We may also assume that X and Y are algebraic sets in some $\mathbb{K}^n$. Then

$$\mathcal{O}_Y(Y) = \mathbb{K}[T_{1|Y}, \ldots, T_{n|Y}] \qquad \mathbb{K}(Y) = \mathbb{K}(T_{1|Y}, \ldots, T_{n|Y}).$$

Let $d = \dim(Y)$. Then the extension $\mathbb{K} \subseteq \mathbb{K}(Y)$ has a transcendence basis of the form $\{T_{i_1|Y}, \ldots, T_{i_d|Y}\}$, because we can choose d algebraically

independent elements among $\{T_1, \ldots, T_n\}$ and these, restricted to Y, will form a basis of the field extension. It follows that if $f \in \mathbb{K}[T_{i_1}, \ldots, T_{i_d}]$ with $f(T_{i_1}, \ldots, T_{i_d})_{|Y} = 0$, then $f = 0$. Hence, the set

$$\{T_{i_1|X}, \ldots, T_{i_d|X}\} \subseteq \mathbb{K}(X)$$

is algebraically independent over $\mathbb{K}$, because for each $f \in \mathbb{K}[T_{i_1}, \ldots, T_{i_d}]$ with $f(T_{i_1}, \ldots, T_{i_d})_{|X} = 0$, we have $f(T_{i_1}, \ldots, T_{i_d})_{|Y} = 0$, which implies $f = 0$. Hence, $\mathrm{trdeg}_{\mathbb{K}}(\mathbb{K}(X)) \geq d$. □

Theorem 2.5.5
Let X be an irreducible affine variety and $Y \subseteq X$ a closed subvariety. If $\dim(Y) = \dim(X)$, then $Y = X$.

Proof. Without loss of generality let X and Y be algebraic sets in $\mathbb{K}^n$, with X irreducible. Let $d = \dim(X) = \dim(Y)$. As in the proof of Theorem 2.5.4, choose a transcendence basis of the form $\{T_{i_1|Y}, \ldots, T_{i_d|Y}\}$ for the extension $\mathbb{K} \subseteq \mathbb{K}(Y)$. As in the previous proof, the set $\{T_{i_1|X}, \ldots, T_{i_d|X}\} \subseteq \mathbb{K}(X)$ is algebraically independent, hence it is a transcendence basis for the extension $\mathbb{K} \subseteq \mathbb{K}(X)$.

By contradiction, assume now $Y \neq X$. Then $I_X(Y) \neq \langle 0 \rangle$ and there is a function $0 \neq f \in \mathcal{O}_X(X)$ such that $f \in I_X(Y)$. The set $\{T_{i_1|X}, \ldots, T_{i_d|X}, f\} \subseteq \mathbb{K}(X)$ is algebraically dependent over $\mathbb{K}$ because it contains $d + 1$ elements. Hence there are $a_j \in \mathbb{K}[T_{i_1}, \ldots, T_{i_d}]$ ($j = 1, \ldots, k$ for some k) and an equation on X

$$a_{k|X} f^k + \ldots + a_{1|X} f + a_{0|X} = 0.$$

Since X is irreducible, $\mathcal{O}_X(X)$ has no proper divisor of zero. Since $f \neq 0$, we may assume $a_{0|X} \neq 0$ (otherwise we may divide by the lowest power of f). Since $f \in I_X(Y)$, we have $a_{0|X} \in I_X(Y)$, that means $a_{0|Y} = 0$. But $a_{0|Y} \in \mathbb{K}[T_{i_1|Y}, \ldots, T_{i_d|Y}]$ and since $a_0 \neq 0$, then $\{T_{i_1|Y}, \ldots, T_{i_d|Y}\}$ are algebraically dependent over $\mathbb{K}$, which is a contradiction. □

To prove the next results, we need to recall the Hilbert's Nullstellensatz.

Recall that for any ideal $\mathfrak{a}$ in a commutative ring R with unity we can define the radical of $\mathfrak{a}$,

$$\sqrt{\mathfrak{a}} = \{r \in R \mid r^n \in \mathfrak{a} \text{ for some } n \in \mathbb{N}\}.$$

This is again an ideal in R.

Theorem 2.5.6 (Hilbert's Nullstellensatz)
Let $\mathbb{K}$ be an algebraically closed field, and $X \subseteq \mathbb{K}^n$ an algebraic set. Then

$$I_X(V_X(\mathfrak{a})) = \sqrt{\mathfrak{a}}$$

for every ideal $\mathfrak{a} \subseteq \mathbb{K}[X]$.

The proof can be found in [ZS58], Chapter VII, § 3, Theorem 14.

Definition 2.5.7
An affine variety X is *purely k-dimensional* if each of its irreducible components has dimension k.

Proposition 2.5.8
Let X be an affine variety, $f_1, \ldots, f_r \in \mathcal{O}_X(X)$ pairwise not associated prime elements and $f = \prod_{i=1}^r f_i$. Then, $V_X(f_i)$ $(i = 1, \ldots, r)$ are the irreducible components of $V_X(f)$.

Proof. We may assume that X and Y are algebraic sets in $\mathbb{K}^n$. By the Corollary 2.5.6 to the Hilbert's Nullstellensatz, $I(Y_i) = \sqrt{\langle f_i \rangle}$. Since each f_i is prime, then $\langle f_i \rangle = \sqrt{\langle f_i \rangle}$ and it follows that each Y_i is irreducible. We have to show that $Y_i \not\subseteq Y_j$ if $i \neq j$. Otherwise, we would have $f_j \in \langle f_i \rangle$ for distinct i, j. But this contradicts the hypothesis that f_i and f_j are not associated. □

Theorem 2.5.9
If $f \in \mathbb{K}[T_1, \ldots, T_n]$ is a non constant polynomial, then the variety $V(f) \subseteq \mathbb{K}^n$ is purely $(n-1)$-dimensional.

Proof. Since $\mathbb{K}[T_1, \ldots, T_n]$ is a unique factorization domain, we can write $f = f_1^{\mu_1} \cdot \ldots \cdot f_r^{\mu_r}$, where the f_is are pairwise not associated. Since $V(f) = V(\prod_{i=1}^r f_i)$, we can apply Proposition 2.5.8, and we obtain that the $V(f_i)$s are the irreducible components of $V(f)$. We have to show that each $V(f_i)$ has dimension $n-1$. Hence, we may assume that f is prime and prove that $V(f)$ has dimension $n-1$.
Let $X = V(f)$ have dimension d. By Theorem 2.5.4, $d \leq n$, and since $f \neq 0$, X is not the whole $\mathbb{K}^n$. By Theorem 2.5.5, $d < n$. Hence, we need only show that $d \geq n-1$. For that purpose, choose a $j \in \{1, \ldots, n\}$ such that f depends on the variable T_j (such a j exists, because f is not constant). Consider the set

$$A = \{T_{1|X}, \ldots, T_{j-1|X}, T_{j+1|X}, \ldots, T_{n|X}\}.$$

We claim that A is algebraically independent over $\mathbb{K}$. Suppose it is not, then there is a polynomial $g \in \mathbb{K}[T_1, \ldots, T_n]$, not depending on T_j, such that

$g_{|X} = 0$. Since $g \in I(X) = \langle f \rangle$, we have $g = hf$ for some polynomial h. This is a contradiction to the fact that g does not depend on T_j.
Hence, $A \subseteq \mathbb{K}(X)$ is algebraically independent over $\mathbb{K}$ and we have $d = \text{trdeg}_{\mathbb{K}}(\mathbb{K}(X)) \geq n-1$. □

The following result considers the inverse situation.

Theorem 2.5.10
If $X \subseteq \mathbb{K}^n$ is a purely $(n-1)$-dimensional algebraic set, then the vanishing ideal $I(X)$ is generated by a polynomial $f \in \mathbb{K}[T_1, \ldots, T_n]$, $I(X) = \langle f \rangle$.

Proof. First, we consider X be irreducible. Then its vanishing ideal $I(X) \subseteq \mathbb{K}[T_1, \ldots, T_n]$ is prime. By Theorem 2.5.5, X is a proper subset of $\mathbb{K}^n$, hence $I(X) \neq \langle 0 \rangle$. Thus, there is an element $0 \neq g \in I(X)$. Since $I(X)$ is a prime ideal, there is a prime factor $f \in I(X)$ of g. Then $X \subseteq V(f)$ and by Theorem 2.5.9, $\dim(V(f)) = n - 1 = \dim(X)$. By Theorem 2.5.5, we get $V(f) = X$. By the Corollary 2.5.6 to the Hilbert's Nullstellensatz, we have $I(X) = \sqrt{\langle f \rangle}$. Since f is prime, $I(X) = \sqrt{\langle f \rangle} = \langle f \rangle$.

Now we are ready for the general case. Suppose $X_1, \ldots, X_r$ are the irreducible components of X. By the first part of the proof, we have $I(X_i) = \langle f_i \rangle$ for some prime elements $f_1, \ldots, f_r$. Since $X_i \neq X_j$ for $i \neq j$, it holds $\langle f_i \rangle \neq \langle f_j \rangle$. Hence the f_i are pairwise not associated.

It is clear that $\langle f_1 \cdot \ldots \cdot f_r \rangle \subseteq I(X)$. Conversely, let $f \in I(X)$. Then $f \in I(X_i) = \langle f_i \rangle$ for each i, and each f_i divides f. Since the f_is are pairwise not associated prime elements, we have that $f_1 \cdot \ldots \cdot f_r$ divides f, that means $f \in \langle f_1 \cdot \ldots \cdot f_r \rangle$.
Hence $I(X) = \langle f_1 \cdot \ldots \cdot f_r \rangle$. □

2.6 Morphisms

Definition 2.6.1
Let $\varphi : X \to Y$ be a morphism between affine varieties.
φ is said to be a *closed embedding* if the image $\varphi(X)$ is closed in Y and φ is an isomorphism between X and $\varphi(X)$.
φ is called *dominant* or *dominating* if $\varphi(X)$ is dense in Y.

Proposition 2.6.2
φ is a closed embedding if and only if the comorphism φ^* is surjective.
φ is dominant if and only if the comorphism φ^* is injective.

A dominant morphism $\varphi : X \to Y$ between irreducible affine varieties defines a monomorphism between the function fields

$$\varphi^* : \mathbb{K}(Y) \to \mathbb{K}(X) \qquad \frac{g}{h} \mapsto \frac{\varphi^*(g)}{\varphi^*(h)}.$$

Definition 2.6.3
φ is said to be *birational* if $\varphi^* : \mathbb{K}(Y) \to \mathbb{K}(X)$ is an isomorphism.

Proposition 2.6.4
If there is a dominant morphism $\varphi : X \to Y$ between affine varieties, then $\dim(X) \geq \dim(Y)$.

Proof. First, we consider the case where both X and Y are irreducible. Since the comorphism is injective, we have a field extension $\mathbb{K}(Y) \cong \varphi^*(\mathbb{K}(Y)) \subseteq \mathbb{K}(X)$. Hence

$$\dim(Y) = \operatorname{trdeg}_{\mathbb{K}}(\mathbb{K}(Y)) \leq \operatorname{trdeg}_{\mathbb{K}}(\mathbb{K}(X)) = \dim(X)$$

Consider the general case, let $X_1, \ldots X_r$ be the irreducible components of X. Then

$$Y = \overline{\varphi(X)} = \overline{\varphi(X_1)} \cup \ldots \cup \overline{\varphi(X_r)}.$$

Each $\varphi(X_i)$ is irreducible, because it is the image of an irreducible space under a continuous map. Thus, also the closure $\overline{\varphi(X_i)}$ are irreducible.
Since the decomposition in irreducible components is unique, the components of Y are some of the $\overline{\varphi(X_i)}$. In particular there is a component $Y_0 = \overline{\varphi(X_j)}$ with maximal dimension. As in the first part of the proof, we obtain

$$\dim(X) \geq \dim(X_i) \geq \dim(Y_0) = \dim(Y).$$

□

Remark 2.6.5
If there is a birational morphism $\varphi : X \to Y$ between affine varieties, then $\dim(X) = \dim(Y)$, because the fields $\mathbb{K}(X)$ and $\mathbb{K}(Y)$ are isomorphic.

The following results will be used in 2.8.4.

Theorem 2.6.6
Let X and Y be irreducible affine varieties with function fields $\mathbb{K}(X)$ and $\mathbb{K}(Y)$.

1. *For any homomorphism $\psi : \mathbb{K}(Y) \to \mathbb{K}(X)$, there is a principal set $X_f \subseteq X$ and a dominant morphism $\varphi : X_f \to Y$ such that $\varphi^* = \psi$.*

2. *For any isomorphism $\psi : \mathbb{K}(Y) \to \mathbb{K}(X)$, there are principal sets $X_f \subseteq X$ and $Y_g \subseteq Y$ and an isomorphism $\varphi : X_f \to Y_g$ such that $\varphi^* = \psi$.*

Let $\varphi : X \to Y$ be a morphism between affine varieties, $x \in X$ and $\varphi(x) = y$. Note that $\varphi^*(\mathfrak{m}_y) \subseteq \mathfrak{m}_x$ and $\varphi^*(\mathfrak{m}_y^2) \subseteq \mathfrak{m}_x^2$, hence $\varphi^*(\mathfrak{m}_y/\mathfrak{m}_y^2) \subseteq \mathfrak{m}_x/\mathfrak{m}_x^2$. Then, for $\delta \in T_{X,x}$, we have $\delta \circ \varphi^* \in T_{Y,y} = \left(\mathfrak{m}_y/\mathfrak{m}_y^2\right)^*$ and we can give the following definition.

The differential of φ in x is the linear mapping

$$\begin{aligned} T_x\varphi : T_{X,x} &\to T_{Y,y} \\ \delta &\mapsto \delta \circ \varphi^*. \end{aligned}$$

Let $\varphi : X \to Y$ and $\psi : Y \to Z$ be morphisms between affine varieties. Let $x \in X$ and $y = \varphi(x)$. Then it is not hard to see that

$$T_x(\psi \circ \varphi) = T_y\psi \circ T_x\varphi$$

and it is obvious that $T_x \,\mathrm{id} = \mathrm{id}$. In particular, if φ is an isomorphism, then also its differential $T_x\varphi$ is an isomorphism for all $x \in X$.

Given a morphism $\varphi : X \to Y$ between affine varieties, $\varphi^*(\mathcal{O}(X))$ is a subring of $\mathcal{O}(X)$, hence $\mathcal{O}(X)$ has a structure of $\varphi^*(\mathcal{O}(X))$-module.

Definition 2.6.7
We say that a morphism $\varphi : X \to Y$ between affine varieties is *finite* if $\mathcal{O}(X)$ is finitely generated over $\varphi^*(\mathcal{O}(Y))$.

Proposition 2.6.8
If $\varphi : X \to Y$ is a finite dominant morphism between irreducible affine varieties, then $\varphi^*(\mathbb{K}(Y)) \subseteq \mathbb{K}(X)$ is an algebraic field extension of finite degree, and $\dim(X) = \dim(Y)$.

Theorem 2.6.9
Let $\varphi : X \to Y$ be a finite morphism of affine varieties and suppose $A \subseteq X$ is a closed subset. Then $\varphi(A)$ is closed in Y and the restriction $\varphi_{|A} : A \to \varphi(A)$ is a finite morphism.

The result 2.6.8 holds even if X and Y are not irreducible, in the following sense.

Proposition 2.6.10
If $\varphi : X \to Y$ is a finite dominant morphism between affine varieties, then $\dim(X) = \dim(Y)$.

Proof. By Proposition 2.6.4, it is sufficient to show that $\dim(X) \leq \dim(Y)$. Choose an irreducible component X' of X with maximal dimension, that means $\dim(X) = \dim(X')$. The image $Y' = \varphi(X')$ is also irreducible, because φ is continuous. By Theorem 2.6.9 the restriction $\varphi_{|X'} : X' \to Y'$ is a finite morphism of affine varieties. By Proposition 2.6.8 we have $\dim(X') = \dim(Y')$. Applying Theorem 2.5.4, we get $\dim(Y') \leq \dim(Y)$, because Y' is a closed subvariety of Y. Hence

$$\dim(X) = \dim(X') = \dim(Y') \leq \dim(Y).$$

□

2.7 Krull dimension

Definition 2.7.1
The *Krull dimension* of a topological space X is the supremum over all integers $r \in \mathbb{N}$ which allow a strictly ascending sequence

$$\emptyset \neq X_0 \subsetneq \ldots \subsetneq X_r \subseteq X$$

of irreducible closed subsets X_i.

Note that this definition is analogous to a characterization of the dimension of vector spaces: in this case maximal chains of vector subspaces are considered to define the dimension of the vector space.

Let $\psi : R \to S$ be a homomorphism of commutative rings with unity. We say that ψ is a *finite homomorphism* if S is a finitely generated module over $\psi(R)$.

Theorem 2.7.2 (Noether's normalization Lemma)
Let R be a finitely generated $\mathbb{K}$-algebra. Then there is an integer $d \in \mathbb{N}$ and a finite monomorphism $\mathbb{K}[T_1, \ldots, T_d] \to R$.

For the proof, we refer to [Shaf77], chapter 2, § 5.

Proposition 2.7.3 (Noether's normalization Lemma, version 2)
Let X be an affine variety. Then there is an integer $d \in \mathbb{N}$ and a finite surjective morphism $X \to \mathbb{K}^d$.

Theorem 2.7.4
The Krull dimension of an affine variety is equal to its dimension.

Proof. First, we prove that the Krull dimension of an affine variety X is bounded by $\dim(X)$. Let $\emptyset \neq X_0 \subsetneq \ldots \subsetneq X_r \subseteq X$ be an ascending sequence of irreducible closed subsets. By Theorem 2.5.4 we have $\dim(X_i) \leq \dim(X)$ and by Theorem 2.5.5 $\dim(X_i) \leq \dim(X_{i+1})$. Then $r \leq \dim(X)$.

Conversely, we have to prove that $\dim(X)$ is bounded by the Krull dimension of X. If $X = \mathbb{K}^n$, we have a sequence $\{0\} \subsetneq \mathbb{K} \subsetneq \mathbb{K}^2 \subsetneq \ldots \subsetneq \mathbb{K}^n$. Then the Krull dimension of $\mathbb{K}^n$ is at least $n = \dim(\mathbb{K}^n)$.

Let X be an arbitrary affine variety. By the Noether's Normalisation Lemma 2.7.3 there is a surjective finite morphism $\varphi : X \to \mathbb{K}^n$, where $n = \dim(X)$.

Let $X_n \subseteq X$ be a n-dimensional irreducible component of X. By Theorem 2.6.9, the morphism

$$\varphi_n = \varphi_{|X_n} : X_n \to \mathbb{K}^n$$

is finite. Let $Y_n = \varphi_n(X_n)$. By Proposition 2.6.8 $\dim(Y_n) = n$. By Theorem 2.5.5 we have $Y_n = \mathbb{K}^n$.

Consider $\mathbb{K}^{n-1} \subsetneq \mathbb{K}^n$; the restriction of φ_n to $\varphi_n^{-1}(\mathbb{K}^{n-1})$ is finite and surjective. Then, $\varphi_n^{-1}(\mathbb{K}^{n-1})$ has dimension $n-1$, by Proposition 2.6.10. In particular there is a $(n-1)$-dimensional irreducible component $X_{n-1} \subseteq \varphi_n^{-1}(\mathbb{K}^{n-1})$. It is $X_{n-1} \subsetneq X_n$ and there is a finite surjective morphism

$$\varphi_{n-1} = \varphi_{n|X_{n-1}} : X_{n-1} \to \mathbb{K}^{n-1}.$$

In the same way we can find a $(n-2)$-dimensional irreducible $X_{n-2} \subsetneq X_{n-1}$, and so on. This proceeding leads to a strictly ascending sequence $X_0 \subsetneq \ldots \subsetneq X_{n-1} \subsetneq X_n$ of irreducible closed sets $X_i \subseteq X$. Since the length of the sequence is n, it follows that the Krull dimension of X is at least n. □

The following theorem extends the result 2.5.9.

Theorem 2.7.5 (Krull's principal ideal Theorem)
Let X be an n-dimensional irreducible affine variety and suppose $f \in \mathcal{O}(X)$ such that $\emptyset \neq V_X(f) \neq X$. Then, $V_X(f)$ is purely $(n-1)$-dimensional.

For the proof, we refer to [ZS58].

Corollary 2.7.6
Let X be an irreducible affine variety and suppose $f_1, \ldots f_r \in \mathcal{O}(X)$ such that $V_X(f_1, \ldots, f_r) \neq \emptyset$. Then, each irreducible component of $V_X(f_1, \ldots, f_r)$ has at least dimension $\dim(X) - r$.

Proof. We use induction over r. The case $r = 1$ is the statement of Krull's Theorem 2.7.5.
For the induction step, consider the subvariety $X' = V_X(f_1, \ldots, f_{r-1})$. We

have $V_X(f_1,\ldots,f_r) = V_{X'}(f_{r|X'})$. For each irreducible component Y of $V_X(f_1,\ldots,f_r)$, choose an irreducible component $Y' \subseteq X'$ such that $Y \subseteq Y'$. By induction hypothesis we have $\dim(Y') \geq \dim(X) - r + 1$. Hence, by Krull's Theorem 2.7.5,

$$\dim(Y) = \dim(V_{Y'}(f_{r|Y'})) \geq \dim(X) - r.$$

□

Corollary 2.7.7
Let X be an irreducible affine variety of dimension n and let x be a point in X. Then there is a strictly ascending chain

$$\{x\} \subsetneq X_1 \subsetneq \ldots \subsetneq X_n,$$

where the X_i are irreducible and closed subsets of X.

Proof. We use induction over $n = \dim(X)$. The case $n = 0$ is trivial. To prove the induction step, define $X_n = X$ and choose a nonzero $f \in \mathcal{O}(X_n)$ with $f(x) = 0$. By Theorem 2.7.5 $V_X(f)$ has an irreducible component X_{n-1} of dimension $n-1$ with $x \in X_{n-1}$. Now apply the induction hypothesis on X_{n-1}. □

2.8 Zariski dimension

Recall that for an irreducible affine variety X, in 2.3.7 we have defined the Zariski dimension as

$$\mathrm{Zdim}\, X = \min_{x \in X} \dim(T_{X,x}).$$

Definition 2.8.1
A polynomial $f \in \mathbb{K}[T]$ is said to be *separable* if each of its irreducible factors has only simple roots in its splitting field.
Let $\mathbb{K} \subseteq \mathbb{L}$ be a field extension. An element $l \in \mathbb{L}$ is *separable over* $\mathbb{K}$ if it is algebraic over $\mathbb{K}$ and its minimum polynomial $f_l \in \mathbb{K}[T]$ is separable.
The field extension $\mathbb{K} \subseteq \mathbb{L}$ is *separable* if it is algebraic and each element $l \in \mathbb{L}$ is separable over $\mathbb{K}$.

If $l_1,\ldots,l_k \in \mathbb{L}$ are separable elements of a field extension $\mathbb{K} \subseteq \mathbb{L}$, then $\mathbb{K} \subseteq \mathbb{K}(l_1,\ldots,l_k)$ is a separable field extension. If $\mathbb{K} \subseteq \mathbb{L}$ and $\mathbb{L} \subseteq \mathbb{M}$ are separable extensions, also $\mathbb{K} \subseteq \mathbb{M}$ is separable.
If $\mathbb{K} \subseteq \mathbb{L}$ is a simple extension, and $\mathbb{L} = \mathbb{K}(l)$, then l is called a *primitive* element for the extension. Recall the following classical result of field theory.

Theorem 2.8.2 (Theorem of the primitive element)
Each finite separable extension $\mathbb{K} \subseteq \mathbb{L}$ is simple, i.e. $\mathbb{L} = \mathbb{K}(l)$ for some primitive element $l \in \mathbb{L}$.

Proposition 2.8.3
Let $\mathbb{K}$ be an algebraically closed field and $\mathbb{K} \subseteq \mathbb{L}$ a finite field extension. Then there is a transcendence basis $\{a_1, \ldots, a_d\} \subseteq \mathbb{L}$ of the extension, such that $\mathbb{K}(a_1, \ldots, a_d) \subseteq \mathbb{L}$ is a separable extension.

For the proof we refer to [ZS58].

Theorem 2.8.4
Let X be an n-dimensional irreducible affine variety. Then, there is an irreducible polynomial $p \in \mathbb{K}[T_1, \ldots, T_{n+1}]$ such that the function field of X is isomorphic to that of $Y = V(p) \subseteq \mathbb{K}^{n+1}$. Furthermore, there exist two isomorphic principal sets $X_0 \subseteq X$ and $Y_0 \subseteq Y$.

Proof. By Proposition 2.8.3 there is a transcendence basis $\{f_1, \ldots, f_n\} \subseteq \mathbb{K}(X)$ for the field extension $\mathbb{K} \subseteq \mathbb{K}(X)$, such that we can decompose $\mathbb{K} \subseteq \mathbb{K}(X)$ in a transcendent extension $\mathbb{K} \subseteq \mathbb{L}$ and in a separable extension $\mathbb{L} \subseteq \mathbb{K}(X)$, where $\mathbb{L} = \mathbb{K}(f_1, \ldots, f_n)$.

Let h be a primitive element for the extension $\mathbb{L} \subseteq \mathbb{K}(X)$, then $\mathbb{K}(X) = \mathbb{L}(h)$. Consider the minimal polynomial $g \in \mathbb{L}[S]$ of h over $\mathbb{L}$. The isomorphism defined by

$$\begin{array}{rccl} \Phi: & \mathbb{L}[S] & \to & \mathbb{K}(T_1, \ldots, T_n)[T_{n+1}] \\ & f_i & \mapsto & T_i \\ & S & \mapsto & T_{n+1} \end{array}$$

maps g to $\Phi(g) = \frac{p}{q}$, where $p \in \mathbb{K}[T_1, \ldots, T_{n+1}]$ and $q \in \mathbb{K}[T_1, \ldots, T_n]$ have no common factor. Since g is irreducible, so is also p. Denote by $Q(R)$ the quotient field of the integral domain R. We have

$$\begin{aligned} \mathbb{K}(V(p)) &= Q(\mathbb{K}[T_1, \ldots, T_{n+1}]/\langle p \rangle) \\ &\cong Q(\mathbb{K}(T_1, \ldots, T_n)[T_{n+1}]/\langle p/q \rangle) \\ &\cong Q(\mathbb{L}[S]/\langle g \rangle) \\ &\cong \mathbb{L}[S]/\langle g \rangle \\ &\cong \mathbb{L}(h) \\ &= \mathbb{K}(X) \end{aligned}$$

We can obtain the isomorphism of principal sets using Theorem 2.6.6. □

Theorem 2.8.5
The Zariski dimension of an irreducible affine variety coincides with its dimension (defined as transcendence degree).

Proof. Let $\dim(X) = n$ be the dimension of the irreducible variety X, defined as transcendence degree. We will show that there is a nonempty open subset $U \subseteq X$ such that $\dim(T_{X,u}) = n$ for each $u \in U$. Then, since the set of regular points X^{reg} is open and dense, the set $X^{\text{reg}} \cap U$ is not empty, hence the dimension of $T_{X,x}$ at each $x \in X^{\text{reg}}$ is n and so $\text{Zdim}(X) = n$.

Let $p \in \mathbb{K}[T_1, \ldots, T_{n+1}]$ be an irreducible polynomial as in Theorem 2.8.4, such that $V(p)$ and X have isomorphic function fields and there are two principal sets $V \subseteq X$ and $W \subseteq V(p)$ and an isomorphism $\varphi : V \xrightarrow{\cong} W$. Since p is irreducible, $I(V(p)) = p$ and for each regular point $y \in V(p)$

$$T_{V(p),y} = V(L_{f,y} \mid f \in I(V(p)) = p) = V(L_{p,y}).$$

Note that $L_{p,y} = \sum_{i=1}^{n+1} \frac{\partial p}{\partial T_i}(y)(T_i - y_i)$ is a non constant polynomial in $n+1$ variables, hence by Theorem 2.5.9, $V(L_{p,y})$ is purely n-dimensional.

Hence we have

$$\dim(T_{V(p),y}) = \dim(V(L_{p,y})) = n.$$

The set $U = \varphi^{-1}(W \cap V(p)^{\text{reg}})$ is open and nonempty in X, and $\dim(T_{X,u}) = n$ for every $u \in U$ and the proof is complete. □

2.9 Krull dimension for rings

The Krull dimension can also be defined on rings.

Definition 2.9.1
The *Krull dimension* of a commutative ring with unity R is the supremum over all $r \in \mathbb{N}$ which allow a strictly ascending sequence

$$\mathfrak{p}_0 \subsetneq \ldots \subsetneq \mathfrak{p}_r \subseteq R$$

of prime ideals.

Proposition 2.9.2
Let X be an irreducible affine variety, and x a point in X. Then

$$\dim(X) = \dim(\mathcal{O}(X)) = \dim(\mathcal{O}_{X,x}).$$

Proof. There is a bijective correspondence between irreducible closed subsets of X and prime ideals in $\mathcal{O}(X)$. In particular if $X_1 \subseteq X_2$ are irreducible closed subsets of X, then $I_X(X_1) \supseteq I_X(X_2)$.

Then the first equality is a direct consequence of Theorem 2.7.4, which says that the Krull dimension of X coincides with $\dim(X)$.

To prove the second equality, observe that the canonical homomorphism $\mathcal{O}_X(X) \to \mathcal{O}_{X,x}$ defines a bijection

$$\begin{array}{rcl} \{\text{Prime ideals in } \mathfrak{m}_x\} & \to & \{\text{Prime ideals in } \mathcal{O}_{X,x}\} \\ \mathfrak{q} & \mapsto & (\mathcal{O}_{X,x} \setminus \mathfrak{m}_{X,x})^{-1}\mathfrak{q}. \end{array}$$

By Corollary 2.7.7 the maximal length of the sequences of prime ideals in $\mathfrak{m}_x$ is the dimension of X. □

Definition 2.9.3
Let R be a local ring with maximal ideal $\mathfrak{m}$. We say that R is regular, if its Krull dimension is equal to the dimension of $\mathfrak{m}/\mathfrak{m}^2$ as vector space over $R/\mathfrak{m}$,

$$\dim(R) = \dim_{R/\mathfrak{m}}(\mathfrak{m}/\mathfrak{m}^2).$$

Theorem 2.9.4
Let X be an irreducible affine variety, and x a point in X. Then x is regular if and only if the local ring $\mathcal{O}_{X,x}$ is regular.

Proof. Suppose x is a regular point, then

$$\dim(\mathcal{O}_{X,x}) = \dim(X) = \dim(T_{X,x}) = \dim(\mathfrak{m}_{X,x}/\mathfrak{m}_{X,x}^2).$$

Conversely, if $\mathcal{O}_{X,x}$ is a regular ring,

$$\dim(T_{X,x}) = \dim(\mathfrak{m}_{X,x}/\mathfrak{m}_{X,x}^2) = \dim(\mathcal{O}_{X,x}) = \dim(X).$$

□

An important result which can be obtained using the theory described in this chapter is that on objects which can be studied both as differentiable manifolds over $\mathbb{C}^n$ and affine varieties on the ground field $\mathbb{C}$, the definitions of dimension given in this and in the first chapter coincide. In particular, one can show that x is a regular point of an affine variety X if and only if it has an open neighbourhood which is diffeomorphic to some $\mathbb{C}^n$. We give a sketch of the proof.

In this chapter we defined the sheaf of regular functions. In the same way, taking holomorphic function instead of regular ones, one can define the sheaf of holomorphic functions of a differentiable manifold. We denote by $\hat{\mathcal{O}}_{X,x}$ the stalk of holomorphic functions at the point $x \in X$.
The following result can be found in [Ser56], §2, Corollary 2.

Theorem 2.9.5
$\mathcal{O}_{X,x}$ has the same dimension as $\hat{\mathcal{O}}_{X,x}$.

Furthermore, defining $\hat{\mathfrak{m}}_{X,x}$ as the maximal ideal $\{h_x \in \hat{\mathcal{O}}_{X,x} \mid h(x) = 0\}$ in $\hat{\mathcal{O}}_{X,x}$, it can be shown that

$$\frac{\mathfrak{m}_{X,x}}{\mathfrak{m}_{X,x}^2} \quad \text{and} \quad \frac{\hat{\mathfrak{m}}_{X,x}}{\hat{\mathfrak{m}}_{X,x}^2}$$

are isomorphic. Hence $\mathcal{O}_{X,x}$ is a regular local ring if and only if $\hat{\mathcal{O}}_{X,x}$ is a regular local ring.

Then, by Theorem 2.9.4, we have that a point $x \in X$ is regular exactly when $\hat{\mathcal{O}}_{X,x}$ is regular.

Suppose x is a smooth point, i.e. there is a biholomorphism from an open neighbourhood U of x, $U \to \mathbb{C}^n$, such that $x \mapsto 0$. Hence $\hat{\mathcal{O}}_{X,x} \cong \hat{\mathcal{O}}_{\mathbb{C}^n,0}$ and the maximal ideal $\mathfrak{m}_x = (x_1, \ldots, x_n)$ has length n, thus $\operatorname{krdim} \hat{\mathcal{O}}_{X,x} = n$.

Conversely, suppose that $x \in X \subseteq \mathbb{C}^n$, where $X = V(I)$ and I is the ideal generated by $f_1, \ldots, f_s$. Suppose the rank of the Jacobi matrix is r, $\operatorname{rk} Df_x = \operatorname{rk}\left(\frac{\partial f_i}{\partial x_j}(x)\right)_{i,j} = r$. Then by the Inverse Function Theorem 1.2.3 there is a system of coordinates for which the following holds:

$$\begin{array}{ll} \text{if } k \leq r & f_k = x_k + g_k(x_{r+1}, \ldots, x_s) \\ \text{if } k > r & f_k = g_k(x_{r+1}, \ldots, x_s) \end{array}$$

In other words, defining the set $X' = \{g_k(x_{r+1}, \ldots, x_s) \mid k > r\}$ and the map

$$\begin{array}{rccl} F: & X' & \to & \mathbb{C}^n \\ & (x_{r+1}, \ldots, x_s) & \mapsto & -f_k(x_{r+1}, \ldots, x_s) \quad k = 1, \ldots, r \end{array}$$

then X is the graph of F. X is biholomorphic to X' and the Jacobi matrix of F is defined by equations with rank 0. Hence we may suppose $r = 0$, that means $I \subseteq \mathfrak{m}_{\mathbb{C}^n,x}^2$.
Then it suffices to show that if $\hat{\mathcal{O}}_{X,x}$ is regular, then $I = (0)$, that means $X = \mathbb{C}^n$.

Since $I \subseteq \mathfrak{m}_{\mathbb{C}^n,x}^2$

$$\frac{\mathfrak{m}_{X,x}}{\mathfrak{m}_{X,x}^2} \cong \left(\frac{\mathfrak{m}_{\mathbb{C}^n,x}}{\mathfrak{m}_{\mathbb{C}^n,x}^2}\right)/I \cong \frac{\mathfrak{m}_{\mathbb{C}^n,x}}{\mathfrak{m}_{\mathbb{C}^n,x}^2}$$

hence $\dim \dfrac{\mathfrak{m}_{X,x}}{\mathfrak{m}_{X,x}^2} = \dim \dfrac{\mathfrak{m}_{\mathbb{C}^n,x}}{\mathfrak{m}_{\mathbb{C}^n,x}^2} = n$.

Thus it suffices to prove that if $I \neq (0)$, then $\operatorname{krdim} \hat{\mathcal{O}}_{X,x} < n$. We do this by contradiction: suppose $X \subsetneq \mathbb{C}^n$ and $\operatorname{krdim} \hat{\mathcal{O}}_{X,x} = n$. Then there is a sequence of prime ideals

$$\mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \ldots \subsetneq \mathfrak{p}_n.$$

Denoting by π the projection

$$\pi : \hat{\mathcal{O}}_{\mathbb{C}^n,x} \twoheadrightarrow \hat{\mathcal{O}}_{X,x}$$

and letting $\mathfrak{q}_i = \pi^{-1}(\mathfrak{p}_i) \subseteq \mathcal{O}_{\mathbb{C}^n,x}$, we find a sequence of prime ideals in $\mathcal{O}_{\mathbb{C}^n,x}$

$$\mathfrak{q}_0 \subsetneq \mathfrak{q}_1 \subsetneq \ldots \subsetneq \mathfrak{q}_n$$

with $(0) \neq I \subseteq \mathfrak{q}_0$. Hence the sequence of prime ideals

$$(0) \subsetneq \mathfrak{q}_0 \subsetneq \mathfrak{q}_1 \subsetneq \ldots \subsetneq \mathfrak{q}_n$$

has length $n+1$, and $\operatorname{krdim} \mathcal{O}_{\mathbb{C}^n,x} > n$, contradiction to the fact that $\mathbb{C}^n$ has dimension n.

Chapter 3

Dimension of metric spaces

In this chapter we study the dimension of topological spaces, in particular of metric spaces. Three different definitions of dimension are given, and it will be proved in Theorems 3.8.2 and 3.8.5 that they coincide on a large class of topological spaces. In Theorem 3.9.3 it will be proved that this notion of dimension coincides with the usual definition for Euclidean spaces, i.e. $\mathbb{R}^n$ has dimension n. From this fact it follows that this dimension extends the concept of dimension given in chapter 1 for manifolds.

3.1 Three concepts of dimension

We first recall some concepts concerning collections of subsets of a space.

Definition 3.1.1
Let $\mathcal{U}$ be a collection of subsets of a topological space R and p a point of R. The *order of* $\mathcal{U}$ *at* p, denoted by $\operatorname{ord}_p \mathcal{U}$ is the number of sets of $\mathcal{U}$ that contain p, if it is finite, and ∞ otherwise.
The *order of* $\mathcal{U}$ is $\operatorname{ord} \mathcal{U} = \sup\{\operatorname{ord}_p \mathcal{U} \mid p \in R\}$, if it is finite, and ∞ otherwise.

Definition 3.1.2
A collection $\mathcal{U}$ of subsets of a space R is said to be *locally finite* if for each point $p \in R$ there is a neighbourhood of p which intersect at most finitely many elements of $\mathcal{U}$.
A collection $\mathcal{U}$ is called *point-finite* if each point is contained in at most finitely many elements of $\mathcal{U}$, i.e. the order of $\mathcal{U}$ is finite.
A *covering* of a space R is a collection $\mathcal{U}$ of subsets of R such that $\bigcup \mathcal{U} = R$.
A *subcovering* $\mathcal{V}$ of $\mathcal{U}$ is a covering with $\mathcal{V} \subseteq \mathcal{U}$.
A *refinement* $\mathcal{V}$ of $\mathcal{U}$ is a covering such that for each $V \in \mathcal{V}$ there is a $U \in \mathcal{U}$ with $V \subseteq U$. Then we write $\mathcal{V} < \mathcal{U}$.

Note that each subcovering is a refinement.

There are three "classical" definitions of dimension.

Definition 3.1.3
The *covering dimension* of a topological space R (also called *Čech-Lebesgue dimension* or *topological dimension*) is denoted by $\dim R$.
We say that $\dim R \leq n$ if for every finite open covering of R there exists an open refinement with order not exceeding $n+1$. Then we define $\dim \emptyset = -1$ and for $R \neq \emptyset$ we set $\dim R = \infty$ if $\dim R > n$ for each integer n and

$$\dim R = \min\{n \mid \dim R \leq n\}$$

otherwise.

Remark 3.1.4
If $\dim R \leq n$ and $\{U_1, \ldots, U_k\}$ is an open covering of R, then there is an open covering $\mathcal{W} = \{W_1, \ldots, W_k\}$ such that

- $\operatorname{ord} \mathcal{W} \leq n+1$ and
- $W_i \subseteq U_i$ for all i.

In fact, by the definition, there is a refinement $\mathcal{V} = \{V_\gamma\}_{\gamma\in\Gamma}$ of $\mathcal{U}$ of order at most $n+1$. Defining

$$\begin{aligned} W_1 &= \bigcup_{V_\gamma \subseteq U_1} V_\gamma \\ W_2 &= \bigcup_{\substack{V_\gamma \subseteq U_i \\ V_\gamma \not\subseteq U_1}} V_\gamma \\ &\vdots \\ W_i &= \bigcup_{\substack{V_\gamma \subseteq U_i \\ V_\gamma \not\subseteq U_1 \cup \ldots \cup U_{i-1}}} V_\gamma \\ &\vdots \end{aligned}$$

we get the desired covering $\{W_1, \ldots, W_k\}$.

Two other definitions of dimension are based on the idea of separating two disjoint closed sets, using an open set, whose boundary has a smaller dimension.

We write ∂A for the boundary of a set A, i.e. $\partial A = \overline{A} \setminus \mathring{A}$.

Definition 3.1.5
The *small* (or *weak*) *inductive dimension* (also called *Urysohn-Menger dimension*) of a topological space R is denoted by $\operatorname{ind} R$ and defined as follows.
For the empty space, set $\operatorname{ind} \emptyset = -1$.
For every non empty space R and for every $n \geq 0$, define $\operatorname{ind} R \leq n$ if for every neighbourhood U of every point p there exists an open neighbourhood V of p such that $p \in V \subseteq U$ and $\operatorname{ind} \partial V \leq n - 1$.
Define

$$\operatorname{ind} R = \min\{n \mid \operatorname{ind} R \leq n\}$$

if this minimum exists, and $\operatorname{ind} R = \infty$ otherwise.

Note that a space has small inductive dimension equal to zero if and only if it is nonempty and has a basis for its topology consisting of open and closed sets. In fact, an open set is closed exactly when its boundary is empty.

It follows that if a space has small inductive dimension equal to zero and its topology is not trivial, then it is not connected.

Definition 3.1.6
The *large* (or *strong*) *inductive dimension* of a topological space R is denoted by $\operatorname{Ind} R$ and defined as follows.
For the empty space, set $\operatorname{Ind} \emptyset = -1$.
For every non empty space R and for every $n \geq 0$, define $\operatorname{Ind} R \leq n$ if for every closed and disjoint subsets F, G of R there exists an open subset $U \subseteq R$ such that $F \subseteq U \subseteq R \setminus G$ and $\operatorname{Ind} \partial U \leq n - 1$.
Define

$$\operatorname{Ind} R = \min\{n \mid \operatorname{Ind} R \leq n\}$$

if this minimum exists, and $\operatorname{Ind} R = \infty$ otherwise.

Note that the properties used to characterize the three dimension are invariant for homeomorphisms, i.e. these dimensions are well defined on topological spaces.

Now, let us look at some examples.

Example 3.1.7
Every nonempty discrete set is zero dimensional.

To show this in the sense of the covering dimension, note that every open covering contains the refinement each of whose elements consists of a single point. This refinement has order 1, hence the covering dimension of the space is at most 0, but the space is not empty, so its dimension is 0.

For the inductive dimensions, note that every subset is open and closed, and thus all the boundaries are empty (their inductive dimension is -1), that is all subsets have inductive dimension 0.

Example 3.1.8
Consider the real line $\mathbb{R}$. Clearly, $\dim \mathbb{R} > 0$, since each refinement of the covering $\{(-\infty, 1), (-1, \infty)\}$ has order at least 2. We show that $\dim \mathbb{R} \leq 1$: let $\mathcal{U}$ be a finite open covering, say containing k open sets $U_1, \ldots, U_k$; each of them is a union of disjoint open intervals

$$U_i = \bigcup_{j \in J_i} (a_i^j, b_i^j) \qquad i = 1, \ldots, k.$$

Consider the set S_0 containing all these intervals,

$$S_0 = \{(a_i^j, b_i^j) \mid j \in J_i, i = 1, \ldots, k\}.$$

Order the rational points $q_1, q_2, \ldots, q_n, \ldots$ and recursively define the sets

$$S_n = S_{n-1} \setminus \left\{(a_i^j, b_i^j) \mid q_n \in (a_i^j, b_i^j)\right\}_{j \in J_i, i=1,\ldots,k} \cup \{(\underline{a}, \underline{b}), (\overline{a}, \overline{b})\},$$

where $(\underline{a}, \underline{b})$ and $(\overline{a}, \overline{b})$ are intervals such that

$$\underline{a} = \min\{a_i^j \mid q_n \in (a_i^j, b_i^j)\} \tag{3.1.1}$$
$$\overline{b} = \max\{b_i^j \mid q_n \in (a_i^j, b_i^j)\}. \tag{3.1.2}$$

Note that such minima and maxima exist, since each point of $\mathbb{R}$ lies in finitely many (more precisely at most k) intervals.

Clearly $S = \bigcap_{n \in \mathbb{N}} S_n$ is an open refinement of $\mathcal{U}$; we show that S has order at most 2, by proving that each point $p \in \mathbb{R}$ is contained in at most two intervals of S. If $p = q_n$ is rational, note that S_n, and hence S, contains at most two intervals which contain q_n, because all the other intervals that contain q_n are not in S_n. If $p \in \mathbb{R} \setminus \mathbb{Q}$ and p lies in three intervals, of whose b is the minimum of the right endpoints, then there is a rational point q such that $p < q < b$, and q is in three intervals, which is a contradiction.

We have to show that S covers $\mathbb{R}$. Since the initial intervals cover $\mathbb{R}$, it is possible that S does not cover $\mathbb{R}$ if and only if all intervals containing some point are not contained in S_n if n is large enough. Hence, it is sufficient to show that when an interval (a, b) is in some S_n but not in S_{n+1}, all points in (a, b) are still contained in some other intervals. To prove this fact, note that $(a, b) \in S_n$ is not contained in S_{n+1} exactly when there are intervals $(\underline{c}, \underline{d})$

and $(\overline{c}, \overline{d})$ such that:

$$\begin{gathered}(\underline{c}, \underline{d}) \cap (\overline{c}, \overline{d}) \neq \emptyset, \\ \underline{c} \leq a, \\ \overline{d} \geq b.\end{gathered}$$

But then $(a, b) \subseteq (\underline{c}, \underline{d}) \cup (\overline{c}, \overline{d})$, hence every point of $\mathbb{R}$ is contained in some intervals of S.

The small inductive dimension of the real line is also equal to 1. Clearly, $\operatorname{ind} \mathbb{R} > 0$, because no open subset of $\mathbb{R}$ has empty boundary. To show that $\operatorname{ind} \mathbb{R} \leq 1$, let p be a point and U a neighbourhood of p. Let $r > 0$ be the distance between p and $\mathbb{R} \setminus U$. Then $p \in S_{r/2}(p) \subset U$ and $\partial S_{r/2}(p)$ consists of two points, which are zero-dimensional in the sense of the small inductive dimension.

Also $\operatorname{Ind} \mathbb{R} = 1$ holds, but we do not prove this result here. It will be clear after section 3.8, where we show that the large inductive dimension and the covering dimension coincide for every metric space. A coincidence theorem holds also for the small inductive dimensions: this dimension coincides with the other two in every metric separable spaces.

In the next sections we prove some result that hold in metric spaces. Some results can be generalized for larger classes of spaces, such as normal spaces, but the concepts of dimension do not generally coincide in such spaces, thus we will not deal with them. The results can be found in the book by Nagata [Nag83].

We will consider mostly the topological and the large inductive dimension, while about the small inductive dimension we will only prove a theorem that states the coincidence of the three definitions of dimension for separable metric spaces.

3.2 Some basic theory

In this section we recall some facts about metrizable spaces and state the Stone's coincidence Theorem and the Nagata-Smirnov Metrization Theorem.

In a metric space R, we denote by $S_r(p)$ the ball of center p and radius r.

Recall that a space R is *metrizable* if it is homeomorphic to a metric space, i.e. if it is possible to define a metric ρ on R such that $\{S_{1/n}(p)\}_{n \in \mathbb{N}}$ is a neighbourhood basis of each $p \in R$. Such a metric is said to be *compatible* with the topology of R.

Definition 3.2.1
Let $\mathcal{U}$ be a collection of subsets of a topological space R.
The *star in a point* p is

$$S(p,\mathcal{U}) = \bigcup \{U \mid p \in U \in \mathcal{U}\}.$$

The *star of a set* A is

$$S(A,\mathcal{U}) = \bigcup \{U \mid U \in \mathcal{U} \text{ and } A \cap U \neq \emptyset\} = \bigcup_{a \in A} S(a,\mathcal{U}).$$

We denote by $\mathcal{U}^{\Delta}$ the family of sets $\{S(p,\mathcal{U}) \mid p \in R\}$ and by $\mathcal{U}^*$ the family $\{S(U,\mathcal{U}) \mid U \in \mathcal{U}\}$.
We write $\mathcal{U}^{\Delta\Delta}$ for $\left(\mathcal{U}^{\Delta}\right)^{\Delta}$ and $\mathcal{U}^{**}$ for $(\mathcal{U}^*)^*$.
A *star refinement* (resp. a *delta refinement*) of a covering $\mathcal{U}$ is a covering such that $\mathcal{V}^*$ (resp. $\mathcal{V}^{\Delta}$) refines $\mathcal{U}$.

Proposition 3.2.2
$\mathcal{U} < \mathcal{U}^{\Delta} < \mathcal{U}^* < \mathcal{U}^{\Delta\Delta}$.

Proof. We need only prove the last assertion, since the other ones are obvious. Let $S(U,\mathcal{U})$ be an element of $\mathcal{U}^*$ and $p \in U$ be a point. For any $U' \in \mathcal{U}$ such that $U \cap U' \neq \emptyset$, since there is a point $p' \in U \cap U'$, we have $U, U' \subseteq S(p',\mathcal{U})$. Hence $p \in S(p',\mathcal{U}) \in \mathcal{U}^{\Delta}$, which implies $U' \subseteq S(p',\mathcal{U}) \subseteq S(p,\mathcal{U}^{\Delta})$. Then, since this holds for any $U' \in \mathcal{U}$ not disjoint from U, we have $S(U,\mathcal{U}) \subseteq S(p,\mathcal{U}^{\Delta}) \in \mathcal{U}^{\Delta\Delta}$, i.e. $\mathcal{U}^* < \mathcal{U}^{\Delta\Delta}$. □

Proposition 3.2.3
Let $\mathcal{U}$ be a locally finite collection. Then

$$\bigcup_{U \in \mathcal{U}} \overline{U} = \overline{\bigcup_{U \in \mathcal{U}} U}.$$

It follows that $\bigcup \mathcal{F}$ is closed if $\mathcal{F}$ is a locally finite closed collection.

Definition 3.2.4
A topological space R is said to be

- T_1, if for every two distinct points p, q there is an open set U such that $p \in U$, $q \notin U$. Note that this is equivalent to require that every set consisting of one single point is closed;
- T_2 or *Hausdorff*, if for every two distinct points p, $q \in R$ there are open disjoint sets U and V such that $p \in U$ and $q \in V$;

- *regular*, if it is T_1 and for every closed F and every point $p \in R \setminus F$ there are disjoint open sets U and V such that $F \subseteq U$ and $q \in V$;
- *normal*, if it is T_1 and for every disjoint closed sets F and G there are disjoint open sets U and V such that $F \subseteq U$ and $G \subseteq V$. It can be shown that the sets U and V can be chosen such that $\overline{U} \cap \overline{V} = \emptyset$;
- *totally normal*, if it is T_1 and every open covering has a star refinement;
- *compact*, if for every open covering there is a finite subcovering;
- *paracompact*, if every open covering has a locally finite open refinement;
- *separable*, if it contains a countable dense subset.

Proposition 3.2.5
Totally normal $\Rightarrow$ *normal* $\Rightarrow$ *regular* $\Rightarrow T_2 \Rightarrow T_1$.

Proof. We need only prove the first implication, since the rest is trivial. Let F and G be disjoint closed subsets of a totally normal space R. Let $\mathcal{U} = \{R \setminus F, R \setminus G\}$. This is an open covering, hence it has a star refinement $\mathcal{V}$. Consider the open sets $U = S(F, \mathcal{V})$ and $V = S(G, \mathcal{V})$. Obviously $F \subseteq U$ and $G \subseteq V$. Seeking for a contradiction, suppose $U \cap V \neq \emptyset$. Then there are $V_1, V_2 \in \mathcal{V}$ such that

$$\begin{gathered} V_1 \in S(F, \mathcal{V}) \quad \text{hence} \quad V_1 \cap F \neq \emptyset \\ V_2 \in S(G, \mathcal{V}) \quad \text{hence} \quad V_2 \cap G \neq \emptyset \\ V_1 \cap V_2 \neq \emptyset \end{gathered}$$

Let $p \in V_1 \cap V_2$. Then $V_1 \cup V_2 \subseteq S(p, \mathcal{V})$. Thus, $S(p, \mathcal{V}) \cap F \neq \emptyset$, which implies $S(p, \mathcal{V}) \not\subseteq R \setminus F$ and in the same way $S(p, \mathcal{V}) \cap G \neq \emptyset$ implies $S(p, \mathcal{V}) \not\subseteq R \setminus G$. This is a contradiction, since $S(p, \mathcal{V}) \in \mathcal{V}^{\Delta} < \mathcal{V}^*$, which is a refinement of $\mathcal{U} = \{R \setminus F, R \setminus G\}$. □

Remark 3.2.6
Every metric space is totally normal, because the balls $\{S_{1/n}(p)\}_{n \in \mathbb{N}}$ form a neighbourhood basis of each point p.

Definition 3.2.7
A collection $\mathcal{U}$ of subsets of a space R is said to be:

- *star finite* (resp. *star countable*) if every member of $\mathcal{U}$ intersect only finitely many (resp. countably many) members of $\mathcal{U}$, i.e. for every $U \in \mathcal{U}$
$$|\{V \in \mathcal{U} \mid U \cap V \neq \emptyset\}| < \infty \text{ (resp. } \leq \aleph_0);$$

- *closure preserving*, if for every subcollection $\mathcal{V}$ of $\mathcal{U}$ the following holds:

$$\overline{\bigcup_{U\in\mathcal{V}} U} = \bigcup_{U\in\mathcal{V}} \overline{U};$$

- *discrete*, if it is closure preserving and any two closures of its members do not intersect.

We state two important results of general topology. They will be useful to prove many results in the next sections, where we consider metric spaces. Here we are not interested in the proofs, for which the reader can refer to the book of Nagata, [Nag83], chapters V and VI.

Theorem 3.2.8 (A.H. Stone's Coincidence Theorem)
A T_2 space is paracompact if and only if it is totally normal.

Corollary 3.2.9
Every metric space is paracompact.

Definition 3.2.10
A collection $\mathcal{U}$ of subsets of a space is said σ-locally finite if it can be written as a countable union of locally finite collections of subsets,

$$\mathcal{U} = \bigcup_{i\in\mathbb{N}} \mathcal{U}_i$$

where each $\mathcal{U}_i$ is locally finite.

Theorem 3.2.11 (Nagata-Smirnov Metrization Theorem)
A regular space is metrizable if and only if it has a σ-locally finite open basis.

3.3 The large inductive dimension

In this and in the following sections we suppose that the spaces are metric. Recall that every metric space is regular, paracompact, and hence normal.

In this section we present some results concerning the large inductive dimension, which we will call dimension for simplicity. In the next three sections we will use the results obtained here to prove some main theorems. The Sum Theorem 3.4.5 links the dimension of a space with the dimension of its closed coverings. The Subspace Theorem 3.5.2 ensures the monotonicity of the dimension: every subset of every space can have at most the same

dimension as the whole space. The Decomposition Theorem 3.5.3 states that each n-dimensional space can be decomposed in $n+1$ zero dimensional subsets. Finally, the Product Theorem 3.6.2 considers the dimension of finite product of spaces.

The following lemma states that a closed subset of a space has dimension not greater than the dimension of the whole space. In Section 3.5 we will prove that this result holds for every subset of the space.

Lemma 3.3.1
Let F be a closed subset of a metric space R. Then $\operatorname{Ind} F \leq \operatorname{Ind} R$.

Proof. We proceed by induction on $\operatorname{Ind} R$.

If $\operatorname{Ind} R = -1$, then $R = \emptyset$, hence $F = \emptyset$.

Assume the statement for the spaces with dimension $\leq n-1$. Let R be a space of dimension $\leq n$ and G, H closed disjoint subsets of F. Then by the definition of Ind, there is an open set $U \subseteq R$ such that $G \subseteq U \subseteq R \setminus H$ and $\operatorname{Ind} \partial U \leq n-1$.

Let $V = U \cap F$. It is open in F and $G \subseteq V \subseteq F \setminus H$. Since $\partial_F V$ is closed in F and also in ∂U (because $\partial_F V = \partial U \cap F$), $\partial_F V \subseteq \partial_F U \subseteq \partial U$. Remember that $\operatorname{Ind} \partial U \leq n-1$. Hence by induction we get $\operatorname{Ind} \partial_F V \leq n-1$. Then $\operatorname{Ind} F \leq n$. □

If two spaces A and B have large inductive dimension n and m respectively, in general it is not true that $\operatorname{Ind}(A \cup B)$ is the maximum or the sum of the dimensions of A and B.

For example, $\operatorname{Ind} \mathbb{Q} = \operatorname{Ind}(\mathbb{R} \setminus \mathbb{Q}) = 0$, but the real line $\mathbb{R}$ has dimension 1. The fact that the sets of rational point and of irrational points are zero dimensional will be proved in Example 3.3.7.

However, the dimension can not "spring" arbitrarily in a union of two spaces, as next Lemma shows for the particular case where one of the spaces is zero dimensional.

Lemma 3.3.2
Let A and B be subsets of a metric space R with $\operatorname{Ind} A \leq n$, $\operatorname{Ind} B \leq 0$ *and* $A \cup B = R$. *Then* $\operatorname{Ind} R \leq n+1$.

To prove the previous Lemma, we need the following result.

Lemma 3.3.3
Let C and D be subsets of a metric space R with $\overline{C} \cap D = \emptyset$ and $C \cap \overline{D} = \emptyset$. Then there are disjoint open sets M and N in R such that $C \subseteq M$ and $D \subseteq N$.

Proof. To each point $p \in C$ assign an $\varepsilon_p \in \mathbb{R}^+$ such that $S_{\varepsilon_p}(p) \cap D = \emptyset$.
To each point $q \in D$ assign an $\varepsilon_q \in \mathbb{R}^+$ such that $S_{\varepsilon_q}(q) \cap C = \emptyset$.
Define

$$M = \bigcup_{p \in C} S_{\frac{1}{2}\varepsilon_p}(p) \quad \text{and} \quad N = \bigcup_{q \in D} S_{\frac{1}{2}\varepsilon_q}(q).$$

□

Proof of Lemma 3.3.2. Let F and G be closed disjoint subsets of R. Since R is normal, there are open sets V, W such that $F \subseteq V$, $G \subseteq W$ and $\overline{V} \cap \overline{W} = \emptyset$.

Without loss of generality, B is not empty (else $A = R$ and all is trivial), hence $\operatorname{Ind} B = 0$, that is for each two closed disjoint subsets of B there is an open and closed set in B which separates them. Hence, there is $U \subseteq B$ open and closed such that $\overline{V} \cap B \subseteq U \subseteq B \setminus \overline{W}$.

We claim that $F \cup U$ and $G \cup (B \setminus U)$ satisfy the assumptions of Lemma 3.3.3, hence there are open disjoint subsets M and N of R such that $F \cup U \subseteq M$ and $G \cup (B \setminus U) \subseteq N$. Then, since U is open and closed and by the property of N, $\partial M \subseteq A$. By Lemma 3.3.1, $\operatorname{Ind} \partial M \leq \operatorname{Ind} A \leq n$.

Now, M is open and $F \subseteq M \subseteq R \setminus G$, $\operatorname{Ind} \partial M \leq n$, hence $\operatorname{Ind} R \leq n+1$.

We have to prove the claim, i.e. that

$$(F \cup U) \cap \overline{G \cup (B \setminus U)} = \emptyset \quad \text{and} \quad \overline{F \cup U} \cap (G \cup (B \setminus U)) = \emptyset$$

It is sufficient to show that the following intersections are empty.

- $F \cap G = \emptyset$ by assumption.
- $F \cap \overline{B \setminus U} \subseteq F \cap \overline{B \setminus (\overline{V} \cap B)} \subseteq V \cap \overline{B \setminus \overline{V}}$ which is contained in $V \cap \overline{B \setminus \overline{V}} \cap B \subseteq (V \setminus U) \cap B = \emptyset$, since $\overline{V} \cap B \subseteq U$.
- $\overline{U} \cap (B \setminus U) = (\overline{U} \cap B) \setminus U$, but $\partial U \subseteq R \setminus B$, since U is open and closed in B, hence $\overline{U} \cap B = U$, hence our intersection is empty.
- This is exactly the same situation as above: since U is open and closed in B, also $B \setminus U$ is so.
- $\overline{U} \cap G \subseteq (\overline{U} \cap W \cap A) \cup (\overline{U} \cap W \cap B)$. The first part is contained in $\overline{B \setminus \overline{W}} \cap W \subseteq [(B \setminus \overline{W}) \cup \partial W] \cap W = [(B \setminus \overline{W}) \cap W] \cup (\partial W \cap W) = \emptyset$. The second part is contained in $\overline{U} \cap (B \setminus U)$, which is empty, as already shown.

□

The following result about point-finite open coverings of a normal space will be useful in the proofs of further results about the large inductive dimension.

Proposition 3.3.4
Let $\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$ be a point-finite open covering of a normal space R. Then there exists an open covering $\mathcal{V} = \{V_\alpha\}_{\alpha\in A}$ such that $\overline{V_\alpha} \subseteq U_\alpha$ for each $\alpha \in A$. (Note that the two coverings have the same cardinality).

Proof. By Zermelo's theorem, we can well-order $\mathcal{U}$ and write $\mathcal{U} = \{U_\alpha\}_{\alpha<\tau}$. We find, for each $\alpha < \tau$, an open set V_α such that

$$R \setminus \left[\bigcup_{\beta<\alpha}\{V_\beta\} \cup \bigcup_{\gamma>\alpha}\{U_\gamma\}\right] \subseteq V_\alpha \subseteq \overline{V_\alpha} \subseteq U_\alpha \tag{3.3.1}$$

To do this, we use transfinite induction.

The set $R \setminus \bigcup_{0<\gamma} U_\gamma$ is closed and contained in U_0, since $\mathcal{U}$ is a covering. Since R is normal, there exists an open set V_0 such that

$$R \setminus \bigcup_{\gamma>0} U_\gamma \subseteq V_0 \subseteq \overline{V_0} \subseteq U_0$$

thus, 3.3.1 holds for $\alpha = 0$.

Now, assume we have defined the open set V_β,satisfying the condition, for every $\beta < \alpha$. We want to define V_α. Let $\mathcal{U}' = \{V_\beta\}_{\beta<\alpha} \cup \{U_\gamma\}_{\gamma\geq\alpha}$.

We claim that $\mathcal{U}'$ is a covering of R. For this purpose, let $p \in R$. If $p \in U_\gamma$ for a $\gamma \geq \alpha$, then clearly $p \in U_\gamma \in \mathcal{U}'$. Else, if $p \notin U_\gamma$ for every $\gamma \geq \alpha$, then, since $\mathcal{U}$ is point-finite, there is a maximal β such that $p \in U_\beta$ (of course $\beta < \alpha$).If $p \in V_{\beta'}$ for a $\beta' < \beta$, then $p \in V_{\beta'} \in \mathcal{U}'$ and we are done. Else, if $p \notin V_{\beta'}$ for each $\beta' < \beta$, then by 3.3.1 (where $\alpha = \beta$)

$$p \in R \setminus \left[\bigcup_{\beta'<\beta} V_{\beta'} \cup \bigcup_{\beta''>\beta} U_{\beta''}\right] \subseteq V_\beta$$

Then $p \in V_\beta \in \mathcal{U}'$ and our claim is proved.

Since $\mathcal{U}'$ is a covering, the closed set $R \setminus [\bigcup_{\beta<\alpha} V_\beta \cup \bigcup_{\gamma>\alpha} U_\gamma]$ is contained in U_α. By the normality of R, there is an open set V_α satisfying 3.3.1.

We only have to prove that $\mathcal{V} = \{V_\alpha\}_{\alpha<\tau}$ is a covering (V_α open and $\overline{V_\alpha} \subseteq U_\alpha$ is clear). We proceed in the same way as for $\mathcal{U}'$: let $p \in R$. Then, since $\mathcal{U}$ is point-finite, there is a maximal $\beta < \tau$ such that $p \in U_\beta$. If $p \notin V_{\beta'}$ for each $\beta' < \beta$, then

$$p \in R \setminus \left[\bigcup_{\beta'<\beta} V_{\beta'} \cup \bigcup_{\beta''>\tau} U_{\beta''}\right] \subseteq V_\beta$$

which proves that $\mathcal{V}$ covers R. □

Lemma 3.3.5
Let R be a metric space with $\operatorname{Ind} R \leq n$. *Then there exists a σ-locally finite open basis $\mathcal{V}$ of R such that* $\operatorname{Ind} \partial V \leq n-1$ *for each $V \in \mathcal{V}$.*

Proof. By the Nagata-Smirnov metrization Theorem 3.2.11, R has a σ-locally finite open basis $\bigcup_{i=1}^{\infty} \mathcal{U}_i$, where each $\mathcal{U}_i$ is a locally finite open covering.

Let $\mathcal{U}_i = \{U_\alpha^i\}_{\alpha \in A_i}$. Then, by Proposition 3.3.4, there is an open covering $\mathcal{W}_i = \{W_\alpha^i\}_{\alpha \in A_i}$ such that $\overline{W_\alpha^i} \subseteq U_\alpha^i$ for each $\alpha \in A_i$.

Since $\operatorname{Ind} R \leq n$, there are open sets V_α^i ($\alpha \in A_i$) such that $\overline{W_\alpha^i} \subseteq V_\alpha^i \subseteq U_\alpha^i$ and $\operatorname{Ind} \partial V_\alpha^i \leq n-1$. Let $\mathcal{V}_i = \{V_\alpha^i\}_{\alpha \in A_i}$. Then $\mathcal{V} = \bigcup_{i=1}^{\infty} \mathcal{V}_i$ is a basis with the requested properties: by construction it is open with $\operatorname{Ind} \partial V \leq n-1$ for each $V \in \mathcal{V}$. $\mathcal{V}$ is σ-locally finite, because the $\mathcal{V}_i$'s are locally finite, since each V_α^i is contained in U_α^i. Finally, $\mathcal{V}$ is a basis, because each $V_\alpha^i \in \mathcal{V}$ is contained in U_α^i, which is an element of a basis; furthermore V_α^i contains W_α^i and the W_α^i's form a covering. □

We consider the inverse situation for $n = 0$.

Lemma 3.3.6
If a metric space R has a σ-locally finite open basis each of whose elements has empty boundary, then $\operatorname{Ind} R \leq 0$.

Proof. Let $\mathcal{V} = \bigcup_{i=1}^{\infty} \mathcal{V}_i$ be a basis with the stated properties, where each $\mathcal{V}_i$ is a locally finite open collection, and $\partial V = \emptyset$ for each $V \in \mathcal{V}$. Let F and G be closed disjoints subsets of R. Then for each i the set

$$U_i = R \setminus \bigcup \{V \mid V \in \bigcup_{j=1}^{i} \mathcal{V}_j \text{ and } V \cap F = \emptyset\}$$

is open and closed (since $\partial V = \emptyset$ and $\cup_{j=1}^{i} \mathcal{V}_j$ is a locally finite collection of open and closed sets).
F is contained in each U_i. Then for the sequence $U_1 \supseteq U_2 \supseteq \ldots \supseteq U_k \supseteq U_{k+1} \supseteq \ldots \supseteq F$ we have $\bigcap_{i=1}^{\infty} U_i = F$. In fact, since $\mathcal{V}$ is a basis, it follows that for each point out of F there is an open neighbourhood of it outside F. In the same way we obtain a sequence of open and closed sets W_i for G such that $W_1 \supseteq W_2 \supseteq \ldots \supseteq W_k \supseteq W_{k+1} \supseteq \ldots \supseteq G$ and $\bigcap_{i=1}^{\infty} W_i = G$.
Let

$$U = \bigcup_{i=1}^{\infty} (U_i \setminus W_i).$$

It is open because each $U_i \setminus W_i$ is open. The set U is also closed: in fact we can write

$$U = (U_1 \setminus W_1) \sqcup \bigsqcup_{i=2}^{\infty} ((U_i \cap W_{i-1}) \setminus W_i)$$

which is a disjoint union of open and closed sets. Now, if a point $x \in R$ is contained in F, then it is also in U, hence in exactly one of the open and closed disjoint sets. Otherwise, if $x \notin F$, then there is a $\bar{k}$ such that $x \notin U_k$ for all $k \geq \bar{k}$, hence $x \in R \setminus U_{\bar{k}}$. If x is contained in none of the disjoint sets whose union is U, then $R \setminus U_{\bar{k}}$ is an open neighbourhood of x that does not intersect U. Conversely, x is contained in one of the open and closed sets, say W, then $(R \setminus U_{\bar{k}}) \cap W$ is an open neighbourhood of x that does not intersect U. Hence U is closed. Furthermore, $F \subseteq U \subseteq R \setminus G$. Since $\partial U = \emptyset$, we have $\operatorname{Ind} R \leq 0$. □

Example 3.3.7
We show that both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are zero dimensional in the sense of the large inductive dimension. We need only show that the dimension is ≤ 0, because the sets are not empty.

Consider first the irrational numbers. Let $q \in \mathbb{Q}$, then the neighbourhood basis $\{S_{1/n}(q)\}_{n\in\mathbb{N}}$ of q has rational endpoints, hence each $S_{1/n}(q) \cap (\mathbb{R} \setminus \mathbb{Q})$ is open in $\mathbb{R} \setminus \mathbb{Q}$ and has empty boundary.

$$\bigcup_{n\in\mathbb{N}, q\in\mathbb{Q}} \{S_{1/n}(q)\}$$

is a σ-locally finite open basis each of whose elements has empty boundary, hence by Lemma 3.3.6 we have $\operatorname{Ind}(\mathbb{R} \setminus \mathbb{Q}) = 0$.

Now we do the same for the set of rational numbers. For each $n \in \mathbb{N}$ find an irrational number a_n such that $\frac{1}{n+1} < a_n < \frac{1}{n}$ and define

$$S'_n(q) = (q - a_n, q + a_n) \subseteq \mathbb{R}.$$

The set of intervals $\{S'_n(q)\}_{n\in\mathbb{N}}$ is clearly a neighbourhood basis of q, each of whose elements has irrational endpoints. Hence, the sets $S'_n(q) \cap \mathbb{Q}$ are open and have empty boundary in $\mathbb{Q}$. Thus,

$$\bigcup_{n\in\mathbb{N}, q\in\mathbb{Q}} \{S'_n(q)\}$$

is a σ-locally finite open basis each of whose elements has empty boundary, hence by Lemma 3.3.6 we have $\operatorname{Ind} \mathbb{Q} = 0$.

The following result will be generalized in Section 3.5 for spaces of finite dimension.

Lemma 3.3.8
Let A be a subset of a metric space R with $\operatorname{Ind} R \leq 0$. *Then* $\operatorname{Ind} A \leq 0$.

Proof. This is a direct consequence of Lemmata 3.3.5 and 3.3.6: from the first lemma follows that there is a σ-locally finite open basis of R each of whose elements has empty boundary. Intersecting each element of the basis with A, we get a basis for A with the same properties, hence $\operatorname{Ind} A \leq 0$. □

3.4 Sum theorem for the large inductive dimension

The aim of this section is to prove that if a space has a locally countable closed covering, its dimension does not exceed the dimension of each closed set in the covering. We do this incrementally.

Lemma 3.4.1
Let $\{F_i\}_{i\in\mathbb{N}}$ be a countable closed covering of a metric space R such that $\operatorname{Ind} F_i \leq 0$ *for each* $i \in \mathbb{N}$. *Then* $\operatorname{Ind} R \leq 0$.

Proof. Let G and H be closed disjoint sets in R. Since $\operatorname{Ind} F_1 \leq 0$, there is an open and closed subset U_1 of F_1 such that $F_1 \cap H \subseteq U_1 \subseteq F_1 \setminus G$. The sets $H \cup U_1$ and $G \cup (F_1 \setminus U_1)$ are closed and disjoint. Since R is normal, there are open sets V_1 and W_1 such that

$$H \cup U_1 \subseteq V_1 \qquad G \cup (F_1 \setminus U_1) \subseteq W_1 \qquad \overline{V_1} \cap \overline{W_1} = \emptyset \tag{3.4.1}$$

Now, since $\operatorname{Ind} F_2 \leq 0$, there is an open and closed subset U_2 of F_2 such that

$$F_2 \cap H \subseteq F_2 \cap \overline{V_1} \subseteq U_2 \subseteq F_2 \setminus \overline{W_1} \subseteq F_2 \setminus W_1 \subseteq F_2 \setminus G$$

The sets $\overline{V_1} \cup U_2$ and $\overline{W_1} \cup (F_2 \setminus U_2)$ are closed and disjoint. Then, by normality of R, there are open sets V_2 and W_2 such that

$$\overline{V_1} \cup U_2 \subseteq V_2 \qquad \overline{W_1} \cup (F_2 \setminus U_2) \subseteq W_2 \qquad \overline{V_2} \cap \overline{W_2} = \emptyset$$

Proceding with these arguments, we get three sequences $\{U_i\}$, $\{V_i\}$ and $\{W_i\}$ ($i \in \mathbb{N}$), where each U_i is an open and closed subset of F_i and all the V_i's and W_i's are open in R, and

$$F_i \cap H \subseteq U_i \subseteq F_i \setminus G \tag{3.4.2}$$

$$\overline{V_{i-1}} \cup U_i \subseteq V_i \qquad \overline{W_{i-1}} \cup (F_i \setminus U_i) \subseteq W_i \tag{3.4.3}$$

$$\overline{V_i} \cap \overline{W_i} = \emptyset \tag{3.4.4}$$

Then, the sets

$$V = \bigcup_{i=1}^{\infty} V_i \qquad W = \bigcup_{i=1}^{\infty} W_i$$

are open and, by virtue of 3.4.1, $H \subseteq V$, $G \subseteq W$. By 3.4.3 and 3.4.4, V and W are disjoint and by 3.4.1 and 3.4.3 $V_i \cup W_i \supseteq U_i \cup (F_i \setminus U_i) = F_i$.

Then, $V \cup W \supseteq \bigcup_{i=1}^{\infty} F_i = R$.

Since V is closed (because $R \setminus V = W$, which is open) and open in R and $H \subseteq V \subseteq R \setminus G$, we have $\operatorname{Ind} R \leq 0$. □

Lemma 3.4.2
Let $\mathcal{F} = \{F_\gamma\}_{\gamma\in\Gamma}$ be a locally finite closed covering of a metric space R such that $\operatorname{Ind} F_\gamma \leq 0$ *for each $\gamma \in \Gamma$. Then* $\operatorname{Ind} R \leq 0$.

Proof. For each $i \in \mathbb{N}$ let $\mathcal{U}_i = \{U_\delta\}_{\delta\in\Delta_i}$ be an open covering such that

$$\operatorname{mesh} \mathcal{U}_i = \sup_{U\in\mathcal{U}_i} \{\operatorname{diam} U\} \ < \ \frac{1}{i}$$

and such that for each $U_\delta \in \mathcal{U}_i$, the closure $\overline{U_\delta}$ intersects finitely many F_γ's. Such a covering exists, since the space is metric and $\mathcal{F}$ is locally finite. We can suppose $\mathcal{U}_i$ is locally finite, since the space is paracompact and the required properties hold also for refinements.

Each $\overline{U_\delta} \cap F_\gamma$ is closed in F_γ, hence by Lemma 3.3.1, $\operatorname{Ind}(\overline{U_\delta} \cap F_\gamma) \leq \operatorname{Ind} F_\gamma \leq 0$. The collection $\{\overline{U_\delta} \cap F_\gamma\}_{\gamma\in\Gamma}$ is a finite closed covering of $\overline{U_\delta}$, thus by Lemma 3.4.1, $\operatorname{Ind} \overline{U_\delta} \leq 0$.

By Proposition 3.3.4, there is an open covering $\mathcal{V}_i = \{V_\delta\}_{\delta\in\Delta_i}$ such that $\overline{V_\delta} \subseteq U_\delta$ for each $\delta \in \Delta_i$. Since $\operatorname{Ind} \overline{U_\delta} = 0$, there is a subset $W_\delta \subseteq \overline{U_\delta}$ which is open and closed in $\overline{U_\delta}$ and such that $\overline{V_\delta} \subseteq W_\delta \subseteq U_\delta$.

Since W_δ is open and closed in $\overline{U_\delta}$, its boundary ∂W_δ is all contained in $R \setminus \overline{U_\delta}$. On the other hand, $W_\delta \subseteq U_\delta$ implies $\overline{W_\delta} \subseteq \overline{U_\delta}$, hence the boundary ∂W_δ is all contained in $\overline{U_\delta}$; hence ∂W_δ is empty, i.e. W_δ is open and closed in R.

The collection $\mathcal{W}_i = \{W_\delta\}_{\delta\in\Delta_i}$ is locally finite, because each $W_\delta \in \mathcal{W}_i$ is contained in $U_\delta \in \mathcal{U}_i$, and $\mathcal{U}_i$ is locally finite. Then, $\mathcal{W} = \bigcup_{i=1}^{\infty} \mathcal{W}_i$ is σ-locally finite. Furthermore, each element in $\mathcal{W}$ is open and closed by the choice of the W_δ. Finally, $\mathcal{W}$ is a basis of R, because $\operatorname{mesh} \mathcal{W}_i \leq \operatorname{mesh} \mathcal{U}_i < \frac{1}{i}$.

Then, by Lemma 3.3.6, $\operatorname{Ind} R \leq 0$. □

Lemma 3.4.3
Let $\{F_i\}_{i\in\mathbb{N}}$ be a countable closed covering of a metric space R such that $\operatorname{Ind} F_i \leq n$ *for each $i \in \mathbb{N}$. Then* $\operatorname{Ind} R \leq n$.

Lemma 3.4.4
Let $\{F_\gamma\}_{\gamma\in\Gamma}$ be a locally finite closed covering of a metric space R such that $\operatorname{Ind} F_\gamma \leq n$ for each $\gamma \in \Gamma$. Then $\operatorname{Ind} R \leq n$.

Proof. We prove the two lemmata in the same way. We proceed by induction on n. If $n = -1$, the assertions are trivially verified.

Assume that the assertions hold for $n-1$. Since $\operatorname{Ind} F_i \leq n$, by Lemma 3.3.5 there is a σ-locally finite open basis $\mathcal{V}_i = \bigcup_{k=1}^{\infty} \mathcal{V}_{i,k}$ of F_i such that

$$\operatorname{Ind} \partial_{F_i} V \leq n-1 \quad \text{for each } V \in \mathcal{V}_i. \tag{3.4.5}$$

Since $\mathcal{V}_i$ is σ-locally finite, so is also $\mathcal{V}_i' = \{\partial_{F_i} V\}_{V\in\mathcal{V}_i}$.

Let $H_i = \bigcup_{V\in\mathcal{V}_i}\{\partial_{F_i} V\}$. Then $\bigcup_{i=1}^{\infty} \mathcal{V}_i'$ is a σ-locally finite closed covering of $\bigcup_{i=1}^{\infty} H_i$. By induction hypothesis and by 3.4.5, we have

$$\operatorname{Ind} \bigcup_{i=1}^{\infty} H_i \leq n-1. \tag{3.4.6}$$

$\mathcal{V}_i$ restricted to $F_i \setminus H_i$ is an open basis of $F_i \setminus H_i$ (it is also σ-locally finite and $\partial_{F_i\setminus H_i} V$ is empty, by definition of H_i). From Lemma 3.3.6 follows $\operatorname{Ind}(F_i \setminus H_i) \leq 0$.

Let $G_i = F_i \setminus \bigcup_{j=1}^{\infty} H_j \subseteq F_i \setminus H_i$. By Lemma 3.3.8, $\operatorname{Ind} G_i \leq 0$. Each G_i is closed in $\bigcup_i G_i$, hence by Lemma 3.4.1 (or Lemma 3.4.2 for the second lemma)

$$\operatorname{Ind} \bigcup_{i=1}^{\infty} G_i \leq 0 \tag{3.4.7}$$

Finally we have $R = \bigcup_{i=1}^{\infty} H_i \cup \bigcup_{i=1}^{\infty} G_i$, hence by 3.4.6, 3.4.7 and Lemma 3.3.2, $\operatorname{Ind} R \leq n$. □

The following is the main result.

Theorem 3.4.5 (Sum Theorem)
Let $\mathcal{F} = \{F_\gamma\}_{\gamma\in\Gamma}$ be a locally countable closed covering of a metric space R such that for each $\gamma \in \Gamma$ holds $\operatorname{Ind} F_\gamma \leq n$. Then $\operatorname{Ind} R \leq n$.

Proof. Since $\mathcal{F}$ is locally countable, there is an open covering $\mathcal{V}$ each of whose elements intersects at most countably many elements of $\mathcal{F}$. Since R is regular and paracompact, there is a locally finite open covering $\mathcal{U} = \{U_\delta\}_{\delta\in\Delta}$ such that $\overline{\mathcal{U}}$ is a refinement of $\mathcal{V}$. Then each $\overline{U_\delta}$ meets at most countably many elements of $\mathcal{F}$. Since $\operatorname{Ind} F_\gamma \leq n$ and $F_\gamma \cap \overline{U_\delta}$ is a closed subset of F_γ, it follows by 3.3.1, that $\operatorname{Ind}(F_\gamma \cap \overline{U_\delta}) \leq n$. Since $\{F_\gamma \cap \overline{U_\delta}\}_{\gamma\in\Gamma}$ covers $\overline{U_\delta}$, by Lemma 3.4.3 we have $\operatorname{Ind} \overline{U_\delta} \leq n$ and by Lemma 3.4.4, since $\overline{\mathcal{U}}$ is a locally finite closed covering, $\operatorname{Ind} R \leq n$. □

3.5 Subspace and decomposition theorems for the large inductive dimension

In Section 3.3 we saw that if R is a metric space with $\operatorname{Ind} R \leq n$, then there exists a σ-locally finite open basis $\mathcal{V}$ of R such that $\operatorname{Ind} \partial V \leq n-1$ for each $V \in \mathcal{V}$. The converse is also true, in fact the following theorem holds.

Theorem 3.5.1
Let $n \geq 0$. For each nonempty metric space R the following holds: $\operatorname{Ind} R \leq n$ *if and only if there exists a σ-locally finite open basis $\mathcal{V}$ of R such that* $\operatorname{Ind} \partial V \leq n-1$ *for each $V \in \mathcal{V}$.*

Proof. We need only prove the "if" part of the statement. For $n = 0$, we have already proved the result in Lemma 3.3.6.
Let $n > 0$. Let $A = \bigcup_{V \in \mathcal{V}} \partial V$ and $B = R \setminus A$. $\{\partial V\}_{V \in \mathcal{V}}$ is a σ-locally finite closed collection. Since $\operatorname{Ind} \partial V \leq n-1$, by the Sum Theorem 3.4.5, $\operatorname{Ind} A \leq n-1$.

$\mathcal{V}$ restricted to B is an open basis of B and satisfies the assumptions of Lemma 3.3.6. Hence $\operatorname{Ind} B \leq 0$. By Lemma 3.3.2, $\operatorname{Ind} R = \operatorname{Ind}(A \cup B) \leq (n-1)+1 = n$. □

Theorem 3.5.2 (Subspace Theorem)
For every subset A of a metric space R, $\operatorname{Ind} A \leq \operatorname{Ind} R$.

Proof. We proceed by induction on the dimension n of the space: the induction basis corresponds to an empty space and is trivially verified.To prove the induction step, let A be a subset of an n dimensional space R and assume the result holds for every $n-1$ dimensional space. By Theorem 3.5.1, there is a σ-locally finite open basis $\mathcal{V}$ of R with $\operatorname{Ind} \partial V \leq n-1$ for each $V \in \mathcal{V}$. Restrict $\mathcal{V}$ to A. Then $\mathcal{V}_{|A}$ is a σ-locally finite open basis of A and $\partial_A V_{|A} \subseteq \partial V$ for each $V \in \mathcal{V}$, hence by induction hypothesis $\operatorname{Ind} \partial_A V_{|A} \leq n-1$ for each $V_{|A} \in \mathcal{V}_{|A}$. By Theorem 3.5.1, we have $\operatorname{Ind} A \leq n$. □

The following result says that a space of dimension n can be decomposed in $n+1$ zero-dimensional subsets.

Theorem 3.5.3 (Decomposition Theorem)
Let $n \geq 0$ and let R be a metric space. Then $\operatorname{Ind} R \leq n$ *if and only if*

$$R = \bigcup_{i=1}^{n+1} A_i,$$

where each A_i is a subset of R such that $\operatorname{Ind} A_i \leq 0$.

Proof. The "if" part was already proved in Lemma 3.3.2. To prove the inverse implication, as in the proof of Theorem 3.5.1, we decompose $R = A_1 \cup B_1$ in two disjoint sets such that $\operatorname{Ind} A_1 \leq 0$ and $\operatorname{Ind} B_1 \leq n-1$. Then we decompose B_1 in $A_2 \cup B_2$, where $\operatorname{Ind} A_2 \leq 0$ and $\operatorname{Ind} B_2 \leq n-2$. Proceeding by induction in this way we get $n+1$ zero-dimensional sets A_i which cover the space. □

Corollary 3.5.4
For every two subsets A, B of a metric space R,

$$\operatorname{Ind}(A \cup B) \leq \operatorname{Ind} A + \operatorname{Ind} B + 1.$$

3.6 Product theorem for the large inductive dimension

In this section we consider the dimension of the product of two spaces and show that it is bounded by the sum of the dimensions of the spaces. We need the following elementary result.

Remark 3.6.1
For every subsets V, W of metric spaces R, S respectively, the boundary of the product subset in the product space is given by the "derivation" formula

$$\partial_{R\times S}(V \times W) = \partial_R(V) \times \overline{W} \ \cup \ \overline{V} \times \partial_S(W).$$

Theorem 3.6.2 (Product Theorem)
Let R, S be metric spaces, not both empty. Then

$$\operatorname{Ind} R \times S \leq \operatorname{Ind} R + \operatorname{Ind} S.$$

Proof. Let $\operatorname{Ind} R = n$, $\operatorname{Ind} S = m$. We proceed by induction on $n+m$.
If one of the two spaces is empty, the assertion is trivially verified.
Let $n, m \geq 0$. By Theorem 3.5.1 there are σ-locally finite open basis $\mathcal{V} = \bigcup_{i=1}^{\infty} \mathcal{V}_i$ and $\mathcal{W} = \bigcup_{i=1}^{\infty} \mathcal{W}_i$ of R and S respectively, where each $\mathcal{V}_i$ and each $\mathcal{W}_i$ is locally finite, such that

$$\operatorname{Ind} \partial_R V \leq n-1 \quad \text{for each } V \in \mathcal{V} \tag{3.6.1}$$
$$\operatorname{Ind} \partial_S W \leq m-1 \quad \text{for each } W \in \mathcal{W} \tag{3.6.2}$$

The collection $\mathcal{V}_i \times \mathcal{W}_i = \{V \times W \mid V \in \mathcal{V}_i, W \in \mathcal{W}_j\}$ is locally finite and open in $R \times S$. Then $\bigcup_{i,j=1}^{\infty} (\mathcal{V}_i \times \mathcal{W}_j)$ is a σ-locally finite open basis of $R \times S$.

By 3.6.1, 3.6.2 and the induction hypothesis, we have: for each $V \in \mathcal{V}$, for each $W \in \mathcal{W}$:

$$\operatorname{Ind}(\partial_R V \times \overline{W}) \leq n - 1 + m$$
$$\operatorname{Ind}(\overline{V} \times \partial_S W) \leq n + m - 1$$

By Remark 3.6.1 and the Sum Theorem 3.4.5, for each $V \times W \in \bigcup_{i,j=1}^{\infty} \mathcal{V}_i \times \mathcal{W}_j$:

$$\operatorname{Ind} \partial_{R\times S}(V \times W) = \operatorname{Ind}\left((\partial_R V \times \overline{W}) \cup (\overline{V} \times \partial_S W)\right) \leq n + m - 1$$

Then by Theorem 3.5.1, $\operatorname{Ind}(R \times S) \leq n + m$. □

3.7 Covering dimension

We turn to consider the covering dimension and prove some results which we will use to prove the Theorems 3.8.2 and 3.8.5 where the coincidence of the different notion of dimension on metric (and separable for ind) spaces is stated.

Remark 3.7.1
For each closed subset F of a space R, we have $\dim F \leq \dim R$.
In fact, setting $\dim R \leq n$, we have to show $\dim F \leq n$. Let $\mathcal{U} = \{U_1, \ldots, U_k\}$ be an open covering of F, i.e. each U_i can be written as $U_i = F \cap V_i$ for an open subset V_i of R. Then $\mathcal{V} = \{V_1, \ldots, V_k, R \setminus F\}$ is an open covering of R and since $\dim R \leq n$, $\mathcal{V}$ has an open refinement $\mathcal{W}$ of order at most $n + 1$. Then $\{W \cap F \mid W \in \mathcal{W}\}$ is a refinement of $\mathcal{U}$, it is open in F and has order at most $n + 1$, hence $\dim F \leq n$.

Note that this result holds more generally for arbitrary subsets of a metric space: in fact, it holds for the large inductive dimension, and in the next section we show that dim and Ind coincide.

The following result proves that in metric spaces it is the same to consider finite or arbitrary coverings in the definition of the covering dimension. In fact, each covering has a locally finite refinement because every metric space is paracompact, and Theorem 3.7.2 shows that one can consider finite or locally finite coverings.

Theorem 3.7.2
For every metric space R, $\dim R \leq n$ *if and only if each open covering of R has a locally finite refinement $\mathcal{V}$ such that* $\operatorname{ord} \mathcal{V} \leq n + 1$.

Proof. The "if" part is trivial, we need only show the inverse implication. Let $\mathcal{U} = \{U_\gamma\}_{\gamma\in\Gamma}$ be an open covering. Since every metric space is paracompact, we can assume $\mathcal{U}$ to be locally finite (otherwise we can take a locally finite refinement). That means, for each point there is a neighbourhood which intersect finitely many elements of $\mathcal{U}$, hence there is an open covering $\mathcal{A}$ such that each $A \in \mathcal{A}$ intersects finitely many elements of $\mathcal{U}$.

Since the space is paracompact, we can find a locally finite open refinement $\mathcal{B}$ of $\mathcal{A}$, and we suppose it to be indexed by an ordinal τ, $\mathcal{B} = \{B_\nu\}_{\nu\in\tau}$. We can assume that, defining $F_\nu = \overline{B_\nu}$ for each ν, $\mathcal{F} = \{F_\nu\}_{\nu<\tau}$ is still a refinement of $\mathcal{A}$ (otherwise shrink the elements of $\mathcal{B}$ as in Lemma 3.3.4). Hence $\mathcal{F}$ is a locally finite closed covering and, since $\mathcal{F}$ refines $\mathcal{A}$,

$$\text{each } F_\nu \text{ meets at most finitely many elements of } \mathcal{U}. \tag{3.7.1}$$

Remark 3.7.1 ensures that

$$\dim F_\nu \leq n \text{ for each } \nu. \tag{3.7.2}$$

Without loss of generality we can assume $F_0 = \emptyset$. Using transfinite induction we construct a sequence of open sets $\{U_{\nu,\gamma}\}_{\gamma\in\Gamma}$ and a sequence of closed sets $\{A_{\nu,\gamma}\}_{\gamma\in\Gamma}$, $\nu < \tau$ such that:

1. $A_{\nu,\gamma} \subseteq F_\nu$;
2. $U_{\nu,\gamma} = U_\gamma \setminus \bigcup_{\beta<\nu} A_{\beta,\gamma}$;
3. $\bigcup_{\gamma\in\Gamma} (U_{\nu,\gamma} \setminus A_{\nu,\gamma}) = R$;
4. ord $\{(F_\nu \cap U_{\nu,\gamma}) \setminus A_{\nu,\gamma}\}_{\gamma\in\Gamma} \leq n+1$.

Setting $A_{0,\gamma} = \emptyset$ and $U_{0,\gamma} = U_\gamma$, the conditions are verified for $\nu = 0$. Let $\alpha < \tau$ and suppose we have already defined $A_{\nu,\gamma}$ and $U_{\nu,\gamma}$ for all $\nu < \alpha$. Define

$$U_{\alpha,\gamma} = U_\gamma \setminus \bigcup_{\nu<\alpha} A_{\nu,\gamma},$$

then the condition 2 is fulfilled. The collection $\{A_{\nu,\gamma}\}_{\nu<\alpha}$ is locally finite, since $A_{\nu,\gamma} \subseteq F_\nu$ and $\mathcal{F}$ is locally finite. Hence, $\bigcup_{\nu<\alpha} A_{\nu,\gamma}$ is closed, as a locally finite union of closed sets. Thus $U_{\alpha,\gamma}$ is open for each γ.

We claim that the collection $\{U_{\alpha,\gamma}\}_{\gamma\in\Gamma}$ is a covering of R.
Suppose by contradiction that there is a point $x \notin U_{\alpha,\gamma}$ for each γ and define

the set $M_x = \{\gamma \in \Gamma \mid x \in U_\gamma\}$. It is a finite set, since $\mathcal{U}$ is locally finite. Since $\mathcal{U}$ is a covering, then x is in some $U_{\bar\gamma}$ for some $\bar\gamma$, thus $x \notin U_{\alpha,\bar\gamma} = U_{\bar\gamma} \setminus \bigcup_{\nu<\alpha} A_{\nu,\bar\gamma}$, hence $x \in A_{\nu(\bar\gamma),\bar\gamma}$ for some $\nu(\bar\gamma) < \alpha$ and for some $\bar\gamma$.
It follows that for each $\gamma \in M_x$ there is an index $\nu(\gamma) < \tau$ such that $\nu(\gamma) < \alpha$ and $x \in A_{\nu(\gamma),\gamma}$. Since M_x is finite, there is an index $\delta < \tau$ such that $\delta < \alpha$ and $\nu(\gamma) \leq \delta$ for each $\gamma \in M_x$.
Now, we distinguish three cases: if $\gamma \in M_x$ such that $\nu(\gamma) = \delta$, then $x \in A_{\delta,\gamma}$.
If $\gamma \in M_x$ such that $\nu(\gamma) < \delta$, then $x \in \bigcup_{\nu<\delta} A_{\nu,\gamma}$, hence $x \notin U_{\delta,\gamma}$.
Finally, if $\gamma \notin M_x$, then $x \notin U_\gamma$.
Hence in every case, $x \notin U_{\delta,\gamma} \setminus A_{\delta,\gamma}$. This is a contradiction to condition 3, and it is proved that $\{U_{\alpha,\gamma}\}_{\gamma\in\Gamma}$ is a covering of R.

Since F_α meets finitely many U_γ and $U_{\alpha,\gamma} \subseteq U_\gamma$, then only finitely many elements of $\{F_\alpha \cap U_{\alpha,\gamma}\}_{\gamma\in\Gamma}$ are nonempty and this collection covers F_α. Since $\dim F_\alpha \leq n$, there is an open covering $\{B_\gamma\}_{\gamma\in\Gamma}$ of F_α such that $B_\gamma \subseteq F_\alpha \cap U_{\alpha,\gamma}$ for each γ and $\operatorname{ord}\{B_\gamma\}_{\gamma\in\Gamma} \leq n+1$. Define

$$A_{\alpha,\gamma} = \overline{(F_\alpha \cap U_{\alpha,\gamma}) \setminus B_\gamma}$$

It is a closed set and $A_{\alpha,\gamma} \subseteq F_\alpha$ for all γ. This verifies condition 1.
Since $(F_\alpha \cap U_{\alpha,\gamma}) \setminus A_{\alpha,\gamma} = B_\gamma$, then $\operatorname{ord}\{(F_\alpha \cap U_{\alpha,\gamma}) \setminus A_{\alpha,\gamma}\}_{\gamma\in\Gamma} \leq n+1$, which verifies 4.
Since $\{B_\gamma\}_{\gamma\in\Gamma}$ covers F_α, then $F_\alpha \subseteq \bigcup_{\gamma\in\Gamma}(U_{\alpha,\gamma} \setminus A_{\alpha,\gamma})$.
Since $\{U_{\alpha,\gamma}\}_{\gamma\in\Gamma}$ covers R and $A_{\alpha,\gamma} \subseteq F_\alpha$ for each γ, then $R \setminus F_\alpha \subseteq \bigcup_{\gamma\in\Gamma}(U_{\alpha,\gamma} \setminus A_{\alpha,\gamma})$, hence $\bigcup_{\gamma\in\Gamma}(U_{\alpha,\gamma} \setminus A_{\alpha,\gamma}) = R$, and condition 3 is verified.

We have now defined the open sets $U_{\alpha,\gamma}$ and the closed sets $A_{\alpha,\gamma}$ for each γ such that the conditions 1, 2, 3 and 4 are verified. The transfinite sequences are thus defined. Let

$$V_\gamma = U_\gamma \setminus \bigcup_{\nu<\tau} A_{\nu,\gamma}.$$

From condition 1 it follows that the closed collection $\{A_{\nu,\gamma}\}_{\nu<\tau}$ is locally finite, hence $\bigcup_{\nu<\tau} A_{\nu,\gamma}$ is closed, hence V_γ is open.

We claim that $\mathcal{V} = \{V_\gamma\}_{\gamma\in\Gamma}$ is a covering of R.
Suppose by contradiction that there is a point $x \notin \bigcup_{\gamma\in\Gamma} V_\gamma$. Then $x \notin U_\gamma \setminus \bigcup_{\nu<\tau} A_{\nu,\gamma}$ for all γ.
Define the set $M_x = \{\gamma \in \Gamma \mid x \in U_\gamma\}$. For each $\gamma \in M_x$ there is a $\nu(\gamma) < \tau$ such that $x \in A_{\nu(\gamma),\gamma}$. Since M_x is a finite set, there is some $\delta < \tau$ such that $\nu(\gamma) \leq \delta$ for all $\gamma \in M_x$.
Now we distinguish three cases: if $\gamma \in M_x$ and $\nu(\gamma) = \delta$, then $x \in A_{\delta,\gamma}$.
If $\gamma \in M_x$ and $\nu(\gamma) < \delta$, then $x \in \bigcup_{\nu<\delta} A_{\nu,\gamma}$, thus $x \notin U_{\delta,\gamma} = U_\delta \setminus \bigcup_{\nu<\delta} A_{\nu,\gamma}$.

If $\gamma \notin M_x$, then $x \notin U_\gamma$.
In every case, $x \notin U_{\delta,\gamma} \setminus A_{\delta,\gamma}$, and this is a contradiction to condition 3. Hence $\mathcal{V}$ is an open covering of R such that $V_\gamma \subseteq U_\gamma$ for each $\gamma \in \Gamma$.

Now we show that ord $\mathcal{V} \leq n+1$.
Note that since by definition

$$F_\nu \cap V_\gamma = F_\nu \cap (U_\gamma \setminus \bigcup_{\nu<\tau} A_{\nu,\gamma}) = (F_\nu \cap U_\gamma) \setminus \bigcup_{\nu<\tau} A_{\nu,\gamma}$$

and by 2

$$(F_\nu \cap U_\gamma) \setminus A_{\nu,\gamma} = (F_\nu \cap (U_\gamma \setminus \bigcup_{\beta<\nu} A_{\beta,\gamma})) \setminus A_{\nu,\gamma} = (F_\nu \cap U_\gamma) \setminus \bigcup_{\beta\leq\nu} A_{\beta,\gamma}$$

then

$$F_\nu \cap V_\gamma \subseteq (F_\nu \cap U_{\nu,\gamma}) \setminus A_{\nu,\gamma}.$$

Hence by condition 4, ord $\{F_\nu \cap V_\gamma\}_{\gamma\in\Gamma} \leq n+1$ for each γ. Since F_ν covers R, then ord $\{V_\gamma\}_{\gamma\in\Gamma} \leq n+1$.

Hence $\mathcal{V}$ is an open refinement of $\mathcal{U}$ with order not exceeding $n+1$ and the proof is complete. □

The following four lemmata consider this situation: take a locally finite collection of open subsets of a space R and for each of its elements U_γ take a closed subset F_γ contained in it. We find sets V_γ that separates each U_γ from the corresponding $R \setminus F_\gamma$, in such a way that the collection of the boundaries ∂V_γ has order not exceeding the dimension of the space. Then we show that instead of the boundaries we can find open sets with the same property.

Finally, we consider countable many collections of this type and show the same result holds for the boundaries.

We will use these results in the proof of the coincidence Theorem 3.8.2 in the next section.

Definition 3.7.3
Two collections $\mathcal{A} = \{A_\gamma\}_{\gamma\in\Gamma}$ and $\mathcal{B} = \{B_\gamma\}_{\gamma\in\Gamma}$ are said to be *similar* if for each finite subset $M \subseteq \Gamma$

$$\bigcap_{\gamma\in M} A_\gamma = \emptyset \quad \Leftrightarrow \quad \bigcap_{\gamma\in M} B_\gamma = \emptyset.$$

For a collection of sets $\mathcal{U} = \{U_\gamma\}_\gamma$, we denote by $\overline{\mathcal{U}}$ the collection $\{\overline{U_\gamma}\}_\gamma$.

Lemma 3.7.4
Let R be a normal space. Let $\{U_\gamma\}_{\gamma\in\Gamma}$ be a locally finite open collection of subsets of R, and $\{F_\gamma\}_{\gamma\in\Gamma}$ a closed collection such that $F_\gamma \subseteq U_\gamma$ for each $\gamma \in \Gamma$. Then there exists an open collection $\mathcal{G} = \{G_\gamma\}_{\gamma\in\Gamma}$ such that

$$F_\gamma \subseteq G_\gamma \subseteq \overline{G_\gamma} \subseteq U_\gamma$$

and $\mathcal{F}$ and $\overline{\mathcal{G}}$ are similar.

Proof. Well order Γ. We use transfinite induction to construct open sets $\{G_\gamma\}_{\gamma\in\Gamma}$ such that $F_\gamma \subseteq G_\gamma \subseteq \overline{G_\gamma} \subseteq U_\gamma$ and for each $\nu \in \Gamma$,

$$\{K_\gamma^\nu\}_{\gamma\in\Gamma} \text{ is similar to } \{F_\gamma\}_{\gamma\in\Gamma} \tag{3.7.3}$$

where

$$K_\gamma^\nu = \begin{cases} \overline{G_\gamma} & \text{if } \gamma \leq \nu \\ F_\gamma & \text{if } \gamma > \nu \end{cases}$$

We can assume $F_0 = \emptyset$, and the initial condition is fulfiled defining $G_0 = \emptyset$. Suppose $\mu \in \Gamma$ such that G_γ is already defined for each $\gamma < \mu$ with the property 3.7.3 for each $\nu < \mu$. Let

$$L_\gamma = \begin{cases} \overline{G_\gamma} & \text{if } \gamma < \mu \\ F_\gamma & \text{if } \gamma \geq \mu \end{cases}$$

Then we show that $\{L_\gamma\}$ is similar to $\{F_\gamma\}$. Suppose $\gamma_1, \ldots, \gamma_r \in \Gamma$. Without loss of generality $\gamma_1 < \ldots < \gamma_j < \mu \leq \gamma_{j+1} < \ldots < \gamma_r$. Then since $L_{\gamma_i} = K_{\gamma_i}^{\gamma_j}$ for all i,

$$\bigcap_{i=1}^r L_{\gamma_i} = \bigcap_{i=1}^r K_{\gamma_i}^{\gamma_j}$$

hence, since $\{K_\gamma^{\gamma_j}\}_{\gamma\in\Gamma}$ is similar to $\{F_\gamma\}_{\gamma\in\Gamma}$ by induction hypothesis,

$$\bigcap_{i=1}^r L_{\gamma_i} = \bigcap_{i=1}^r K_{\gamma_i}^{\gamma_j} = \emptyset \Leftrightarrow \bigcap_{i=1}^r F_{\gamma_i} = \emptyset$$

Since $L_\gamma \subseteq U_\gamma$ for each γ, the collection $\{L_\gamma\}$ is locally finite. Denote by Λ the set of finite subsets of Γ. For each $\lambda \in \Lambda$ let $E_\lambda = \bigcap_{\gamma\in\lambda} L_\gamma$. Each E_λ is a finite intersection of closed sets. Thus, the collection $\{E_\lambda\}_{\lambda\in\Lambda}$ is closed and locally finite.

Let $E = \bigcup\{E_\lambda \mid E_\Lambda \cap F_\mu = \emptyset\}$ be the union of all finite intersections of L_γ's which do not intersect F_μ. It is closed (because union of locally finite closed sets) and $E \cap F_\mu = \emptyset$.

Since the space is normal, we can separate the closed sets F_μ and $(R \setminus U_\mu) \cup E$, i.e. we can find an open set G_μ such that

$$F_\mu \subseteq G_\mu \subseteq \overline{G_\mu} \subseteq U_\mu \setminus E.$$

In particular, $\overline{G_\mu} \cap E = \emptyset$.

Now the open sets G_γ are defined for $\gamma \leq \mu$. We need only show that $\{K_\gamma^\mu\}$ is similar to $\{L_\gamma\}$, which is similar to $\{F_\gamma\}$. For this purpose, let $\gamma_1, \ldots, \gamma_r \in \Gamma$.

If $\bigcap_{i=1}^r K_{\gamma_i}^\mu = \emptyset$, then $\bigcap_{i=1}^r L_{\gamma_i} = \emptyset$ since $L_{\gamma_i} \subseteq K_{\gamma_i}^\mu$ for each i.

Conversely, suppose $\bigcap_{i=1}^r L_{\gamma_i} = \emptyset$. We have to show that $\bigcap_{i=1}^r K_{\gamma_i}^\mu = \emptyset$. Without loss of generality $\gamma_1 < \ldots < \gamma_j \leq \mu < \gamma_{j+1} < \ldots < \gamma_r$. If $\gamma_j \neq \mu$, we are done, because then $K_{\gamma_i}^\mu = L_{\gamma_i}$ for all i. Suppose $\gamma_j = \mu$, then

$$\emptyset = \bigcap_{i=1}^r L_{\gamma_i} = L_{\gamma_1} \cap \ldots \cap L_{\gamma_{j-1}} \cap F_\mu \cap L_{\gamma_{j+1}} \cap \ldots \cap L_{\gamma_r}. \tag{3.7.4}$$

If the set

$$\bigcap_{i=1}^r K_{\gamma_i}^\mu = L_{\gamma_1} \cap \ldots \cap L_{\gamma_{j-1}} \cap \overline{G_\mu} \cap L_{\gamma_{j+1}} \cap \ldots \cap L_{\gamma_r} \tag{3.7.5}$$

is not empty, there is a point $p \in \overline{G_\mu}$. Hence $p \notin E$, and by construction of E, $p \notin L_\gamma$ for all γ such that $L_\gamma \cap F_\mu = \emptyset$. Since by 3.7.5, $p \in L_{\gamma_i}$ for $i = 1, \ldots, j-1, j+1, \ldots, r$, we have $L_{\gamma_i} \cap F_\mu \neq \emptyset$, hence $\bigcap_{i=1}^r L_{\gamma_i} \neq \emptyset$, which contradicts 3.7.4. □

Lemma 3.7.5
Let R be a metric space with $\dim R \leq n$. *Let $\{U_\gamma\}_{\gamma\in\Gamma}$ be a locally finite open collection of subsets of R, and $\{F_\gamma\}_{\gamma\in\Gamma}$ a closed collection such that $F_\gamma \subseteq U_\gamma$ for each $\gamma \in \Gamma$. Then there exists an open collection $\{V_\gamma\}_{\gamma\in\Gamma}$ such that*

$$F_\gamma \subseteq V_\gamma \subseteq \overline{V_\gamma} \subseteq U_\gamma$$

and ord $\{\partial V_\gamma\}_{\gamma\in\Gamma} \leq n$.

Proof. For each γ define the open covering $\mathcal{W}_\gamma = \{U_\gamma, R \setminus F_\gamma\}$. Define the covering

$$\bigwedge_{\gamma\in\Gamma} \mathcal{W}_\gamma = \left\{ \bigcap_{\gamma\in\Gamma} W_\gamma \mid W_\gamma \in \mathcal{W}_\gamma \text{ for each } \gamma \in \Gamma \right\}$$

Note that each element of $\wedge\{\mathcal{W}_\gamma\}$ is a finite intersection of U_γ's intersected with all other $R \setminus F_\gamma$. In other words, denoting by Λ the set of finite subsets of Γ,

$$\bigwedge_{\gamma\in\Gamma} \mathcal{W}_\gamma = \left\{ \bigcap_{\gamma\in\lambda} U_\gamma \cap \bigcap_{\gamma\notin\lambda} (R \setminus F_\gamma) \mid \lambda \in \Lambda \right\}.$$

Since the collection $\{U_\gamma\}$ is locally finite, then $\wedge\{\mathcal{W}_\gamma\}$ is locally finite, too. Each of its elements is open, because the U_γ's in the intersection are finitely many, and $\cap(R \setminus F_\gamma)$ is open, since F_γ is locally finite. Hence $\wedge\{\mathcal{W}_\gamma\}$ is a locally finite open covering.

By Theorem 3.7.2 there is a locally finite open covering $\mathcal{N} = \{N_\delta\}_{\delta\in\Delta}$ which refines $\wedge\{\mathcal{W}_\gamma\}_{\gamma\in\Gamma}$ and with $\operatorname{ord}\mathcal{N} \leq n+1$. Each N_δ meets finitely many F_γ's, because $\mathcal{N}$ refines $\wedge\{\mathcal{W}_\gamma\}$ and $\{U_\gamma\}$ is locally finite. By Proposition 3.3.4, there is a closed covering $\{K_\delta\}_{\delta\in\Delta}$ with $K_\delta \subseteq N_\delta$ for each δ.

To each N_δ and each F_γ intersecting N_δ assign an open set $N_\delta(F_\gamma)$ such that:

$$K_\delta \subseteq N_\delta(F_\gamma) \subseteq \overline{N_\delta(F_\gamma)} \subseteq N_\delta \text{ and} \tag{3.7.6}$$

$$\overline{N_\delta(F_\gamma)} \subseteq N_\delta(F_{\gamma'}) \text{ (or viceversa) for } \gamma \neq \gamma'. \tag{3.7.7}$$

This is possible because there are only finitely many F_γ's which intersect N_δ. The sets $N_\delta(F_\gamma)$ are related to the sets F_γ only for the cardinality.

Let

$$V_\gamma = \bigcup_{\delta\in\Delta} N_\delta(F_\gamma).$$

Note that in the union we can consider only the δ's such that N_δ intersects F_γ, since the other $N_\delta(F_\gamma)$'s are empty. For each γ, let D_γ be the set of indices $\delta \in \Delta$ such that $F_\gamma \cap N_\delta \neq \emptyset$, then we can write $V_\gamma = \bigcup_{\delta\in D_\gamma} N_\delta(F_\gamma)$. We claim that

$$F_\gamma \subseteq V_\gamma \subseteq \overline{V_\gamma} \subseteq U_\gamma.$$

To show the first inclusion, note that $\{K_\delta\}_{\delta\in D_\gamma}$ is a covering of F_γ, hence

$$F_\gamma \subseteq \bigcup_{\delta\in D_\gamma} K_\delta \subseteq \bigcup_{\delta\in D_\gamma} N_\delta(F_\gamma) = V_\gamma.$$

To show the last inclusion, note that since each N_δ intersects finitely many F_γ's, say $F_{\gamma_1}, \ldots, F_{\gamma_k}$, $N_\delta \subseteq U_{\gamma_1} \cap \ldots \cap U_{\gamma_k} \cap \bigcap_{\gamma\neq\gamma_i} (R \setminus F_\gamma)$. Hence $\overline{N_\delta(F_\gamma)} \subseteq N_\delta \subseteq U_\gamma$. Since $\overline{V_\gamma}$ is a locally finite union of $\overline{N_\delta(F_\gamma)}$'s,

$$\overline{V_\gamma} = \overline{\bigcup_{\delta\in D_\gamma} N_\delta(F_\gamma)} = \bigcup_{\delta\in D_\gamma} \overline{N_\delta(F_\gamma)} \subseteq \bigcup_{\delta\in D_\gamma} N_\delta$$

it follows that $\overline{V_\gamma} \subseteq U_\gamma$.

We now prove that $\{V_\gamma\}$ has the required properties. Clearly, it is an open collection, and we have just seen that $F_\gamma \subseteq V_\gamma \subseteq \overline{V_\gamma} \subseteq U_\gamma$. It remains to prove that ord $\{\partial V_\gamma\} \leq n$.

By contradiction, suppose that ord $\{\partial V_\gamma\}_{\gamma\in\Gamma} > n$, that means we can find distinct γ_i's $(i = 1, \ldots, n+1)$ such that:

$$\bigcap_{i=1}^{n+1} \partial V_\gamma \neq \emptyset$$

Since each V_γ is a locally finite union of some $N_\delta(F_\gamma)$, there are some $N_{\delta_i}(F_{\gamma_i})$ $(i = 1, \ldots, n+1)$ such that:

$$\bigcap_{i=1}^{n+1} \left(\overline{N_{\delta_i}(F_{\gamma_i})} \setminus N_{\delta_i}(F_{\gamma_i})\right) \neq \emptyset \tag{3.7.8}$$

If $\delta_i = \delta_j$ for $i \neq j$, from 3.7.7 we obtain

$$\left(\overline{N_{\delta_i}(F_{\gamma_i})} \setminus N_{\delta_i}(F_{\gamma_i})\right) \cap \left(\overline{N_{\delta_j}(F_{\gamma_j})} \setminus N_{\delta_j}(F_{\gamma_j})\right) = \emptyset$$

which is a contradiction to 3.7.8. Hence the δ_i's $(i = 1, \ldots, n+1)$ must be all distinct.

Let

$$p \in \bigcap_{i=1}^{n+1} \left(\overline{N_{\delta_i}(F_{\gamma_i})} \setminus N_{\delta_i}(F_{\gamma_i})\right) \tag{3.7.9}$$

From 3.7.6 it follows that $p \notin \bigcup_{i=1}^{n+1} K_{\delta_i}$.

Since $\{K_\delta\}$ is a covering of R, there is a δ with $p \in K_\delta$. Thus, $\delta \neq \delta_i$ for all $i = 1, \ldots, n+1$.

By 3.7.6 and 3.7.9 it follows that $p \in \bigcap_{i=1}^{n+1} N_{\delta_i}$, hence

$$p \in K_\delta \cap \bigcap_{i=1}^{n+1} N_{\delta_i} \subseteq N_\delta \cap \bigcap_{i=1}^{n+1} N_{\delta_i}$$

which is a contradiction to ord $\mathcal{N} \leq n+1$. Then, ord $\{\partial V_\gamma\}_{\gamma\in\Gamma} \leq n$. □

Lemma 3.7.6

Let R be a metric space with $\dim R \leq n$. *Let $\{U_\gamma\}_{\gamma\in\Gamma}$ be a locally finite open collection of subsets of R, and $\{F_\gamma\}_{\gamma\in\Gamma}$ a closed collection such that $F_\gamma \subseteq U_\gamma$ for each $\gamma \in \Gamma$. Then there exist open collections $\{V_\gamma\}_{\gamma\in\Gamma}$ and $\{W_\gamma\}_{\gamma\in\Gamma}$ such that*

$$F_\gamma \subseteq V_\gamma \subseteq \overline{V_\gamma} \subseteq W_\gamma \subseteq U_\gamma$$

and ord $\left\{\overline{W_\gamma} \setminus V_\gamma\right\}_{\gamma\in\Gamma} \leq n$.

Proof. By Lemma 3.7.5 there is an open collection $\mathcal{V} = \{V_\gamma\}_{\gamma\in\Gamma}$ such that $F_\gamma \subseteq V_\gamma \subseteq \overline{V_\gamma} \subseteq U_\gamma$ and ord $\{\partial V_\gamma\}_{\gamma\in\Gamma} \leq n$.

Since the space is normal and $\mathcal{U}$ is locally finite, by Lemma 3.7.4 there is an open collection $\mathcal{H} = \{H_\gamma\}_{\gamma\in\Gamma}$ with $\partial V_\gamma \subseteq H_\gamma \subseteq \overline{H_\gamma} \subseteq U_\gamma$ and $\{\partial V_\gamma\}$ is similar to $\overline{\mathcal{H}}$.

For each $\gamma \in \Gamma$ define $W_\gamma = V_\gamma \cup H_\gamma$. Then

$$\partial V_\gamma \subseteq \overline{W_\gamma} \setminus V_\gamma \subseteq \overline{H_\gamma}$$

Hence, since $\{\partial V_\gamma\}$ is similar to $\{\overline{H_\gamma}\}$, it is also similar to $\{\overline{W_\gamma} \setminus V_\gamma\}$. Hence, ord $\{\overline{W_\gamma} \setminus V_\gamma\} \leq n$.

Further, the following inclusions hold:

$$F_\gamma \subseteq V_\gamma \subseteq \overline{V_\gamma} \subseteq W_\gamma \subseteq U_\gamma$$

□

Lemma 3.7.7
Let R be a metric space with $\dim R \leq n$. *Let $\{U_\gamma\}_{\gamma\in\Gamma_i}$, $i \in \mathbb{N}$, be locally finite open collections of subsets of R, and $\{F_\gamma\}_{\gamma\in\Gamma_i}$ closed collections such that $F_\gamma \subseteq U_\gamma$ for each $\gamma \in \Gamma_i$, for each $i \in \mathbb{N}$. Then there exist open subsets V_γ ($\gamma \in \bigcup_{i\in\mathbb{N}} \Gamma_i$) such that*

$$F_\gamma \subseteq V_\gamma \subseteq \overline{V_\gamma} \subseteq U_\gamma$$

and ord $\{\partial V_\gamma\}_{\gamma\in\bigcup_{i\in\mathbb{N}}\Gamma_i} \leq n$.

Proof. Suppose the sets of indices are disjoint. Let $\gamma \in \Gamma_1$, then by Lemma 3.7.6 there are open collections $\mathcal{V}_1 = \{V_\gamma^1\}_{\gamma\in\Gamma_1}$ and $\mathcal{W}_1 = \{W_\gamma^1\}_{\gamma\in\Gamma_1}$ such that

$$F_\gamma \subseteq V_\gamma^1 \subseteq \overline{V_\gamma^1} \subseteq W_\gamma^1 \subseteq U_\gamma$$

and ord $\{\overline{W_\gamma^1} \setminus V_\gamma^1\}_{\gamma\in\Gamma_1} \leq n$.

Now let $\gamma \in \Gamma_1 \cup \Gamma_2$, then by Lemma 3.7.6 there are open collections $\mathcal{V}_2 = \{V_\gamma^2\}_{\gamma\in\Gamma_1\cup\Gamma_2}$ and $\mathcal{W}_2 = \{W_\gamma^2\}_{\gamma\in\Gamma_1\cup\Gamma_2}$ such that

$$\begin{aligned} \overline{V_\gamma^1} \subseteq V_\gamma^2 \subseteq \overline{V_\gamma^2} \subseteq W_\gamma^2 \subseteq W_\gamma^1 \quad &\text{if } \gamma \in \Gamma_1 \\ F_\gamma \subseteq V_\gamma^2 \subseteq \overline{V_\gamma^2} \subseteq W_\gamma^2 \subseteq U_\gamma \quad &\text{if } \gamma \in \Gamma_2 \end{aligned}$$

and ord $\{\overline{W_\gamma^2} \setminus V_\gamma^2\}_{\gamma\in\Gamma_1\cup\Gamma_2} \leq n$.

Proceeding in this way, we get a sequence of open collections $\mathcal{V}_i = \{V_\gamma^i\}_{\gamma\in\Gamma_1\cup\ldots\Gamma_i}$ and $\mathcal{W}_i = \{W_\gamma^i\}_{\gamma\in\Gamma_1\cup\ldots\Gamma_i}$ such that

$$F_\gamma \subseteq \overline{V_\gamma^i} \subseteq W_\gamma^i \subseteq U_\gamma \tag{3.7.10}$$

$$\operatorname{ord}\{\overline{W_\gamma^i} \setminus V_\gamma^i\}_{\gamma\in\Gamma_1\cup\ldots\cup\Gamma_i} \leq n \tag{3.7.11}$$

$$\overline{V_\gamma^i} \subseteq V_\gamma^{i+1} \tag{3.7.12}$$

$$\overline{W_\gamma^i} \supseteq W_\gamma^{i+1} \tag{3.7.13}$$

For $\gamma \in \Gamma_i$ define

$$V_\gamma = \bigcup_{i=1}^{\infty} V_\gamma^i$$

Then, since for each integer j, $F_\gamma \subseteq V_\gamma^j$ and $\overline{V_\gamma^j} \subseteq U_\gamma$, we have

$$F_\gamma \subseteq V_\gamma \subseteq \overline{V_\gamma} \subseteq U_\gamma$$

Let $\gamma \in \Gamma_i$. Then

$$\overline{V_\gamma} \setminus V_\gamma = \bigcup_{j=1}^{\infty} \overline{V_\gamma^j} \setminus \bigcup_{j=1}^{\infty} V_\gamma^j \subseteq \bigcap_{j=1}^{\infty} \left(\overline{V_\gamma^j} \setminus V_\gamma^j\right) \subseteq \bigcap_{j=1}^{\infty} \left(\overline{W_\gamma^j} \setminus V_\gamma^j\right)$$

which, for each fixed integer k, is contained in $\overline{W_\gamma^k} \setminus V_\gamma^k$. The collection $\{\overline{W_\gamma^k} \setminus V_\gamma^k\}_\gamma$ has order not exceeding n, thus it holds $\operatorname{ord}\{\partial V_\gamma\}_\gamma \leq n$. □

3.8 Coincidence theorems

Now we are ready to prove that in metric spaces the dimensions dim and Ind coincide.

Lemma 3.8.1
Let $n \geq 0$. Let $\mathcal{V}$ be a σ-locally finite open basis of a metric space R, such that $\operatorname{ord}\{\partial V\}_{V\in\mathcal{V}} \leq n$. *Then* $\operatorname{Ind} R \leq n$.

Proof. We proceed by induction on n: the case $n = 0$ was already proved in Lemma 3.3.6. Let $n > 0$ and suppose the result holds for $n-1$. Let $\mathcal{V} = \{V_\gamma\}_{\gamma\in\Gamma}$. Then $\mathcal{V}_\gamma = \{\partial V_\gamma \cap V_{\gamma'}\}_{\gamma'\in\Gamma}$ is a σ-locally finite open basis of ∂V_γ with $\operatorname{ord}\{\partial_{\partial V_\gamma} V \mid V \in \mathcal{V}_\gamma\} \leq n-1$, because every point of ∂V_γ lies in at most other $n-1$ boundaries, since the order of ∂V is at most n.
By induction hypothesis, $\operatorname{Ind}\partial V_\gamma \leq n-1$; thus, by Theorem 3.5.1, $\operatorname{Ind} R \leq n$. □

Theorem 3.8.2 (Coincidence Theorem I)
For every metric space R, $\dim R = \operatorname{Ind} R$.

Proof. First, let $\operatorname{Ind} R \leq n$; we show $\dim R \leq n$.
By the decomposition Theorem 3.5.3, we can write $R = \bigcup_{i=1}^{n+1} A_i$, where $\operatorname{Ind} A_i \leq 0$ for each i. Let $\mathcal{U} = \{U_j\}_{j=1,\dots,k}$ be an open covering of R. We claim that for each i there is an open covering $\{V_j\}_{j=1,\dots,k}$ of A_i such that

$$\operatorname{ord} \{V_j\}_{j=1,\dots,k} \leq 1 \qquad V_j \subseteq U_j \tag{3.8.1}$$

To see that such a covering exists, fix an index i and let $Z_j = U_j \cap A_i$. Then $\{Z_j\}_j$ covers A_i. By Proposition 3.3.4 there is a covering $\{Y_j\}_j$ of A_i such that $Y_j \subseteq \overline{Y_j} \subseteq Z_j$. Since $\operatorname{Ind} A_i \leq 0$, we can separate each $\overline{Y_j}$ from $A_i \setminus \overline{Z_j}$ with an open and closed set X_j, that means $\overline{Y_j} \subseteq X_j \subseteq Z_j$. The collection $\{X_j\}_j$ is a covering of A_i, because $\{\overline{Y_j}\}_j$, is a covering and each X_j contains $\overline{Y_j}$.
Now set $V_j^i = X_j \setminus \bigcup_{k<j} X_j$ for each j. The V_j's are open and closed, pairwise disjoint and cover A_i, furthermore $V_j \subseteq U_j$, that means this is a covering with the requested properties.

For each point $p \in V_j$ choose $\varepsilon(p) > 0$ such that

$$S_{\varepsilon(p)}(p) \cap A_i \subseteq V_j$$
$$S_{\varepsilon(p)}(p) \subseteq U_j$$

and define

$$W_j = \bigcup_{p \in V_j} \{S_{\frac{1}{2}\varepsilon(p)}(p)\}.$$

Since $W_j \cap A_i = V_j$, then $\mathcal{W}_i = \{W_j\}_{j=1,\dots,k}$ is an open covering of A_i, $W_j \subseteq U_j$ for each j.

We want show that ord $\mathcal{W}_i \leq 1$. By contradiction, suppose there is a point $p \in W_{j_1} \cap W_{j_2}$ for distinct indices $0 \leq j_1, j_2 \leq k$. Then by the definition of the W_j's, there are $x \in V_{j_1}$ and $y \in V_{j_2}$ such that

$$\begin{array}{lll} p \in S_{\varepsilon(x)/2}(x) & S_{\varepsilon(x)}(x) \subseteq U_{j_1} & S_{\varepsilon(x)}(x) \cap A_i \subseteq V_{j_1} \\ p \in S_{\varepsilon(y)/2}(y) & S_{\varepsilon(y)}(y) \subseteq U_{j_2} & S_{\varepsilon(y)}(y) \cap A_i \subseteq V_{j_2}. \end{array}$$

It follows that

$$d(x,y) \leq d(x,p) + d(p,y) < \frac{\varepsilon(x)}{2} + \frac{\varepsilon(y)}{2} \leq \max\{\varepsilon(x), \varepsilon(y)\},$$

that means $x \in S_{\varepsilon(y)}(y)$ or $y \in S_{\varepsilon(x)}(x)$. Since these conditions are symmetric, suppose the first holds. Then, $x \in V_{j_1} \subseteq A_i$ implies $x \in S_{\varepsilon(y)}(y) \cap A_i \subseteq V_{j_2}$.

Hence, $x \in V_{j_1} \cap V_{j_2}$, which contradicts 3.8.1. Thus, we have shown ord $\mathcal{W}_i \leq 1$.

Define $\mathcal{W} = \bigcup_{i=1}^{n+1} \mathcal{W}_i$. This is an open refinement of $\mathcal{U}$ with order $\leq n+1$, thus $\dim R \leq n$.

Now, to complete the proof, let $\dim R \leq n$; we have to show $\operatorname{Ind} R \leq n$. Since R is paracompact, we can take a sequence $\mathcal{U}_i = \{U_\gamma\}_{\gamma \in \Gamma_i}$, $i \in \mathbb{N}$, of locally finite open coverings such that $\lim_{i\to\infty} \operatorname{mesh} \mathcal{U}_i = \sup_{U\in\mathcal{U}_i}\{\operatorname{diam} U\} = 0$, as in the proof of Lemma 3.4.2. Using Proposition 3.3.4 we can construct closed coverings $\mathcal{F}_i = \{F_\gamma\}_{\gamma\in\Gamma_i}$, $i \in \mathbb{N}$, with $F_\gamma \subseteq U_\gamma$ for all γ. By Lemma 3.7.7 we can find open sets V_γ, $\gamma \in \bigcup_{i=1}^{\infty} \Gamma_i$ such that $F_\gamma \subseteq V_\gamma \subseteq \overline{V_\gamma} \subseteq U_\gamma$ and ord $\{\partial V_\gamma\}_{\gamma\in\bigcup\Gamma_i} \leq n$. Then $\mathcal{V} = \{V_\gamma\}_{\gamma\in\bigcup\Gamma_i}$ is a σ-locally finite open collection. That it is a basis can be seen using the same argument as in the proof of Lemma 3.4.2. Hence by Lemma 3.8.1, $\operatorname{Ind} R \leq n$. □

The Coincidence Theorem 3.8.2 can fail outside the class of metric spaces. We report an example of a compact Hausdorff space S such that

$$\dim S = 1 \quad \text{and} \quad \operatorname{ind} S = \operatorname{Ind} S = 2.$$

We describe briefly how to construct the space S, without proving the assertions. The description of this example and all the proofs can be found in [Pea75], chapter 4, §3.4.

Example 3.8.3
Let Y_1 and Y_2 be two spaces, each of which is homeomorphic to the long line. Identify the last element of Y_1 with the last element of Y_2, and call this point y_0. Define the set

$$S = \{(y,t) \in Y \times I \mid y = y_0 \text{ or } t \in C\},$$

where C denotes the Cantor set. For $i = 1, 2$, let

$$S_i = \{(y,t) \in S \mid y \in Y_i\}.$$

Write each point of C in its ternary expansion, i.e. $\sum_{i=1}^{\infty} \frac{t_n}{3^n}$, where each t_n is 0 or 2. Let $D \subseteq C$ consist of the points in C such that writing the ternary expansion, $t_n = 0$ almost always or $t_n = 2$ almost always. Then D is dense in C. Define a map

$$\begin{array}{rccc} \psi_1 : & C & \to & I \\ & \displaystyle\sum_{i=1}^{\infty} \frac{t_n}{3^n} & \mapsto & \displaystyle\sum_{i=1}^{\infty} \frac{t'_n}{2^n} \end{array}$$

where

$$t'_n = \begin{cases} 0 & \text{if } t_n = 0 \\ 1 & \text{if } t_n = 2. \end{cases}$$

This map is continuous, order-preserving and surjective. Let $M_1 = \psi_1(D) \setminus \{0,1\}$; it is the set of dyadic rationals in the unit interval. In particular, M_1 is countable and dense in the unit interval.
The function

$$\begin{array}{rccc} f: & I & \to & I \\ & x & \mapsto & \dfrac{e^x - 1}{e - 1} \end{array}$$

is surjective and strictly ascending. We claim that f maps each rational in $(0,1)$ on an irrational number. In fact, suppose $x = \frac{p}{q}$ is a rational number in $(0,1)$ such that $f(x) = r$, where r is rational. Then

$$e^p = (r(e-1)+1)^q$$

i.e. e is solution of a polynomial equation, which is impossible, since e is transcendent.
Define

$$\begin{array}{rccc} \psi_2: & C & \to & I \\ & x & \mapsto & f \circ \psi_1(x) \end{array}$$

and let $M_2 = \psi_2(D) \setminus \{0,1\}$. Then M_2 is countable, and it is a dense subset of the open unit interval.
Note that M_1 and M_2 are disjoint. The maps $\psi_i : C \to I$, $i = 1, 2$ are continuous, order-preserving and surjective.
We say that two points $t, s \in D$,

$$t = \sum_{i=1}^{\infty} \frac{t_n}{3^n} \qquad s = \sum_{i=1}^{\infty} \frac{s_n}{3^n}$$

are related if there exists a positive integer m such that $s_n = t_n$ for all $n < m$, $s_m = 0$, $t_m = 2$ and $s_n = 2$, $t_n = 0$ for all $n > m$ (or the same holds, exchanging s with t). It is easy to see that D is partitioned into related pairs. If $s, t \in C$, then $\psi_1(s) = \psi_1(t)$ if and only if s and t are related points in D. Considering C as the unit interval where open intervals are deleted, then D consists of the endpoints of the deleted intervals, and a pair of points is related exactly when they are the endpoints of some deleted interval.
Let Y_1 and Y_2 have the interval topology, and define

$$g_i : Y_i \times C \to S_i$$

setting

$$g_i(y,t) = \begin{cases} (y,t) & \text{if } y < y_0 \\ (y_0, \psi_i(t)) & \text{if } y = y_0 \end{cases}$$

Let S_i have the identification topology with respect to the surjection g_i, i.e. we take the largest topology on S_i such that g_i is continuous.
Since S_i is the continuous image of the Hasdorff compact space $Y_i \times C$, S_i is compact and it can be proved that it is also Hausdorff.
It can be shown (Proposition 4.3.2 [Pea75]) that $\dim S_i = \operatorname{Ind} S_i = \operatorname{ind} S_i = 1$.
The set $S_1 \cap S_2 = \{y_0\} \times I$ is closed both in S_1 and S_2 and has its usual topology both as a subspace of S_1 and of S_2. Hence we can give S the weak topology with respect to the covering $\{S_1, S_2\}$, i.e. the largest topology on S that induces the given topology on S_1 and on S_2.
Then S is a compact Hausdorff space and S_1 and S_2 are closed in S. It can be shown that, since $S = S_1 \cup S_2$ and the S_i are closed,

$$\dim S = 1.$$

Furthermore, one can prove that

$$\operatorname{ind} S = \operatorname{Ind} S = 2.$$

Remark 3.8.4
Every separable metric space R has a countable basis.
In fact, let $A \subseteq R$ be a countable dense subset and define

$$\mathcal{B} = \left\{ B_{1/2^n}(x) \mid x \in A, n \in \mathbb{N} \right\}.$$

It is easy to see that $\mathcal{B}$ is a countable basis of R.

Theorem 3.8.5 (Coincidence Theorem II)
For every separable metric space R, $\dim R = \operatorname{Ind} R = \operatorname{ind} R$.

Proof. We need only prove $\operatorname{Ind} R \leq \operatorname{ind} R$, since the equivalence of Ind and dim was already proved in the preceding Theorem and $\operatorname{ind} \leq \operatorname{Ind}$ is clear from the definitions of these concepts of dimension. We proceed by induction on $\operatorname{ind} R$. The induction basis, that is the case $\operatorname{ind} R = -1$, is trivial.

Assume $\operatorname{Ind} S \leq \operatorname{ind} S$ for each separable space S with $\operatorname{ind} S \leq n-1$ and let $\operatorname{ind} R = n$, where R is a metric separable space, hence by Remark 3.8.4, it has a countable open basis $\{U_i\}_{i \in \mathbb{N}}$. We define open sets W_{ij}, $i, j \in \mathbb{N}$ in the following way:

- if there exists an open set W such that $\overline{U_i} \subseteq W \subseteq U_j$ with $\operatorname{ind} \partial W \leq n-1$ (and $\operatorname{Ind} \partial W \leq n-1$ by induction hypothesis), then set $W_{ij} = W$;
- otherwise $W_{ij} = \emptyset$.

Since $\operatorname{ind} R = n$, $\{W_{ij}\}_{i,j\in\mathbb{N}}$ is an open basis with $\operatorname{Ind} \partial W_{ij} \leq \operatorname{ind} \partial W_{ij} \leq n-1$. By Theorem 3.5.1 $\operatorname{Ind} R \leq n$. □

The hypothesis that the space is separable is necessary: an example of a metrizable space P such that

$$\operatorname{ind} P = 0 \quad \text{and} \quad \dim P = \operatorname{Ind} P = 1$$

was given by Prabir Roy in [Roy62] and [Roy68], and it is described also in [Pea75], chapter 7 §4.

3.9 The dimension of the Euclidean space

In this section we show that the dimension of the Euclidean space $\mathbb{R}^n$ is n. To prove it, we introduce the notion of geometrical complex and prove the Sperner's Lemma, which we will use to prove that $\dim \mathbb{R}^n \geq n$.

Definition 3.9.1
The *(geometrical) n-simplex* with vertices $p_0, \ldots, p_n \in \mathbb{R}^m$ is:

$$[p_0, \ldots, p_n] = \left\{ \sum_{i=0}^{n} \lambda_i p_i \mid \sum_{i=0}^{n} \lambda_i = 1, 0 \leq \lambda_i \leq 1 \right\}.$$

The *k-face* $[p_{i_0}, \ldots, p_{i_k}]$ is the set of points of the simplex such that $\lambda_i = 0$ for each $i \neq i_j$, where $j = 0, \ldots, k$.
A finite set K of geometrical simplexes in $\mathbb{R}^m$ is a (*geometrical*) *complex* if

- every face of each simplex in K belongs to K and
- the intersection of two simplexes in K is either a face of both simplexes or it is empty.

A triangulation of a space X consists of a complex K together with a homeomorphism $h : K \to X$, i.e. it is a subdivision of the space into simplices.

Lemma 3.9.2 (Sperner's Lemma)
Let $\{S_i\}_{i=0,\ldots,n}$ be the $(n-1)$-faces of an n-simplex T^n and $\{F_i\}_{i=0,\ldots,n}$ a covering of T^n such that $F_i \subseteq T^n \setminus S_i$ for each $i = 0, \ldots, n$. Let K be a complex which triangulates T^n. Then there is an n-simplex of K which intersects every F_i.

Proof. To each vertex p of K assign an integer $i(p)$ such that $p \in F_{i(p)}$. To each r-simplex $T = [p_0, \ldots, p_r]$ of K assign the set $i(T) = \{i(p_0), \ldots, i(p_r)\}$. Let $\{T_j\}_{j=1,\ldots,k}$ be the n-simplexes of K and s the number of T_j's with $i(T_j) = \{0, \ldots, n\}$.

We have to show $s > 0$. To do that, it is enough to prove that s is odd.

We proceed by induction on n. If $n = 0$, the claim is trivially true.

Assume the claim for the $(n-1)$-simplexes. Let s_j be the number of $(n-1)$-dimensional faces T' of T_j such that $i(T') = \{1, \ldots, n\}$. For each j it must be $s_j = 0, 1$ or 2. For, if $s_j \neq 0$, there is a face $T' = [p_1, \ldots, p_n]$ of $T_j = [p_0, \ldots, p_n]$ such that $i(T') = \{1, \ldots, n\}$. Now, if $i(p_0) = 0$, T' is the only face with the requested property, hence $s_j = 1$. Else, $1 \leq i(p_0) \leq n$ implies $s_j = 2$.

Let t be the number of T_j's such that $s_j = 2$. Then

$$\sum_{j=1}^{k} s_j = s + 2t \tag{3.9.1}$$

because the term s comes from the T_j's with $s_j = 1$ and the term $2t$ from the T_j's with $s_j = 2$.

Let u be the number of $(n-1)$-simplexes T' of K such that they are face of exactly one of the T_j's, i.e.

$$\begin{cases} i(T') = \{1, \ldots, n\} \\ T' \subseteq \bigcup_{i=0}^{n} S_i \end{cases} \tag{3.9.2}$$

Let v be the number of $(n-1)$-simplexes T'' of K such that they are face of two of the T_j's, i.e.

$$\begin{cases} i(T'') = \{1, \ldots, n\} \\ T'' \not\subseteq \bigcup_{i=0}^{n} S_i \end{cases} \tag{3.9.3}$$

Then, $\sum_{j=1}^{k} s_j = u + 2v$.

Combining this with 3.9.1, we get $u + 2v = s + 2t$.

It is sufficient to show that u is odd. If T' is an $(n-1)$-simplex with 3.9.2, $T' \subseteq S_0$ (because if $T' \subseteq S_i$ for an $i \neq 0$, since $F_i \subseteq T^n \setminus S_i$, then $i \notin i(T')$, contradiction).

Therefore, u is the number of $(n-1)$-simplexes T' of K contained in S_0 and such that $i(T') = \{1, \ldots, n\}$. Applying the induction hypothesis to the covering $\{F_i \cap S_0\}_{i=1,\ldots,n}$ of the $(n-1)$-simplex S_0, we get that u is odd. □

Theorem 3.9.3
$\dim \mathbb{R}^n = n$.

Proof. The covering dimension of $\mathbb{R}$ is ≤ 1, as we saw in Example 3.1.8. The Product Theorem 3.6.2 together with the Concidence Theorem 3.8.2 ensures that $\dim \mathbb{R}^n \leq n$.

To show that $\dim \mathbb{R}^n \geq n$, we consider a simplex $T^n \subset \mathbb{R}^n$. By Remark 3.7.1, we have $\dim T^n \leq \dim \mathbb{R}^n$. Then, it is enough to prove that T^n has dimension at least n.

Let $\{S_i\}_{i=0,\dots,n}$ be the $(n-1)$-faces of T^n. Consider the open covering $\{T^n \setminus S_i\}_{i=0,\dots,n}$ of T^n. By Remark 3.1.4, we can find an open covering $\mathcal{V} = \{V_i\}_{i=0,\dots,n}$ of T^n such that $V_i \subseteq T^n \setminus S_i$ for each i and $\operatorname{ord} \mathcal{V} \leq \dim T^n + 1$. Then there is a closed covering $\{F_i\}_{i=0,\dots,n}$ of T^n with $F_i \subseteq V_i$.

We have to prove that $\cap_{i=0}^n F_i \neq \emptyset$. Then, $\cap_{i=0}^n V_i \neq \emptyset$, hence $\operatorname{ord} \mathcal{V} \geq n+1$, hence $\dim T^n \geq n$.

By contradiction, suppose $\cap_{i=0}^n F_i = \emptyset$, then since T^n is compact, $\{T^n \setminus F_i\}_{i=0,\dots,n}$ is an uniform covering of T^n, that means, there is an $\varepsilon > 0$ such that $\{S_\varepsilon(p) \mid p \in T^n\}$ is a refinement. Here ε is called a Lebesgue number for the space.

Now construct a triangulation K of T^n, such that each of its simplexes has diameter at most ε. Then each simplex of K is contained in $T^n \setminus F_i$ for some i, contradicting Lemma 3.9.2. □

Corollary 3.9.4
Each open subset of $\mathbb{R}^n$ has dimension n.

In fact, each open subset of $\mathbb{R}^n$ contains a n-simplex as subset.

Hence, every n-dimensional manifold (where the dimension is the one defined in chapter 1) has (covering) dimension n, because the dimension of manifolds is given via local homeomorphisms with open subsets of the Euclidean spaces. In other words, the concept of dimension for metric spaces studied in this chapter extends the concept of dimension for manifolds, studied in chapter 1.

As a last note, we investigate how the definitions of topological dimensions apply to algebraic sets with the Zariski topology.

Remark 3.9.5
The covering dimension of an algebraic set counts its cardinality if it is finite, and otherwise it is not defined.

Proof. Let X be an irreducible algebraic set (or an irreducible component of an algebraic set) and take n distinct points $x_1, \dots, x_n \in X$. Define the (Zariski-)open sets $U_i = X \setminus \{x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$, for $1 \leq i \leq n$. Then $\mathcal{U} = \{U_i\}_{1 \leq i \leq n}$ is a covering and each of its refinements contains at

least n elements. Remember that every finite intersection of (Zariski-)open sets is not empty. Therefore every refinement of $\mathcal{U}$ has order at least n, and the covering dimension of X can be made arbitrarily large, except in the case where X is a finite set of points, where the covering dimension is the cardinality of X. □

Remark 3.9.6
The inductive dimensions on algebraic sets coincide with the Krull dimension.

Proof. We proceed by induction on the Krull dimension of the algebraic set X. If X is empty, it has inductive dimensions and Krull dimension -1.
Suppose $\operatorname{ind} X \leq \operatorname{krdim} X$ for every algebraic set X with $\operatorname{krdim} X \leq n-1$.
Let X be an irreducible algebraic set such that $\operatorname{krdim}(X) = n$ and let Y be a (Zariski-)closed set in X. Hence $\operatorname{krdim} Y \leq n-1$. Take a point $p \in X \setminus Y$ in case we consider the small inductive dimension or a closed set $Y' \subseteq X$ disjoint from Y in case consider the large inductive dimension. Consider the open set $U = X \setminus Y$. It contains p or Y'. Note that $\partial U = \overline{U} \setminus U = Y$, since every open set is dense in X. Hence $\operatorname{ind} \partial U \leq n-1$ by induction hypothesis and it follows that $\operatorname{ind} X \leq \operatorname{krdim} X$.
To show that the dimensions coincide, let Y be an irreducible closed set in X with $\operatorname{krdim} Y = n-1$ (it exists by Theorem 2.5.9), and let $p \in X \setminus Y$ (or $Y' \subseteq X \setminus Y$). Then for each open neighbourhood U of p disjoint from Y (or each open set $U \supseteq Y'$ disjoint from Y), Y is contained in the closed set $X \setminus U$, which has dimension at least $n-1$, since it contains the closed set Y of dimension $n-1$. Furthermore, $\partial U = X \setminus U$ since X is irreducible and U is dense in X. That means $\operatorname{ind} X = n = \operatorname{krdim} X$. □

Part II

Informal learning of mathematics

Chapter 4

An exhibition about the 4D space

The interuniversitary research centre for the communication and informal learning of mathematics *matematita*[1] is preparing an exhibition about the four dimensional Euclidean space. The exhibition is planned to be ready in the next future. Nevertheless a part of it has already been set up and "tested" in Genova, within the "Festival della Scienza" (Science Festival, 25th October–6th November 2007). The work we are presenting here is part of this project; we attended to the realization of one of the exhibits and of the related interactive animations. In this chapter we describe the exhibition which was set up in Genova giving particular attention to the section we worked for.

The title of the exhibition is "Un tuffo nella quarta dimensione" ("A Trip into the Fourth Dimension"). It is thought for non specialist visitors with a middle or high education level and for schools. In the exhibition some 4D objects are described and represented through 3D models, and problems about their properties are posed.

A deep comprehension of the topics is not expected from the visitor, who is mainly invited to reflect on the concept of dimension and on some facts that characterize the different "usual" dimensions (i.e. the lower dimensions 1, 2 and 3). Dimensional analogy is then used to explain what happens "jumping" from the three dimensional to the four dimensional space and comparisons with the the passage from the 2D space to the 3D space are often made.

[1] `http://www.matematita.it/`

4.1 Some basic definitions

We introduce the definitions of the figures and concepts that are described informally in the exhibition.

A hyperplane in an n-dimensional space is an affine subspace of codimension 1. It is the set of zeroes of a linear equation $a_1x_1+a_2x_2+\ldots+a_nx_n = a_0$ where $a_i \in \mathbb{R}$ $(0 \leq i \leq n)$.

A *polytope* is a compact convex region of an n-dimensional space enclosed by a finite number of hyperplanes, with nonempty interior. The last condition means that we consider only n-dimensional polytopes in a space of dimension n. Polytopes in dimension 0, 1, 2 and 3 are points, segments, polygons and polyhedra respectively.

A *face* is the intersection of an n-dimensional polytope π_n with a tangent hyperplane, i.e. a hyperplane that does not intersect the interior of the polytope. Every face of π_n is a polytope π_k for some $k \in \{0, 1, \ldots, n-1\}$ and is therefore called k-face. A 0-face is called vertex, a 1-face is called edge, and an $(n-1)$-face is called facet.

If the mid-points of all the edges that emanate from a given vertex O of a n-dimensional polytope π_n lie on a hyperplane, then they are the vertices of a $(n-1)$-dimensional polytope, called the vertex figure of π_n at O.

A polytope π_n $(n > 2)$ is said to be *regular* if its facets are regular and there is a regular vertex figure at every vertex. Then it can be shown that the facets are all equal and the vertex figures are all equal. The k-skeleton of an n-polytope $(0 \leq k \leq n)$ consists of all faces of dimension up to k.

An *n-cube*, also called measure polytope, is the product of n intervals of equal length. It is an n-dimensional regular polytope, each of whose facets is parallel to exactly one other facet (opposite to it) and orthogonal to each other facet. A 2-cube is called square, a 3-cube simply cube, and we call a 4-cube hypercube.

Informally speaking, an *unfolding* of a polyhedron is obtained by cutting its surface along some edges and flattening it into the plane, in such a way that the plane figure (also called net) is connected. The unfolding is not unique. The procedure of folding up a net in dimension 3 allows to fold the figure along the edges. In general one has to specify which edges of the net are to be joined, since there might be ambiguity, but for the polyhedra we consider the folding procedure is unique.

The same can be made in higher dimensions: an unfolding of an n-polytope is a $(n-1)$-dimensional connected diagram consituted by the facets of the polytope. It can be folded up along the $(n-2)$-faces to obtain the original polytope in dimension n.

For example, a net of a four dimensional polytope is formed by polyhedra connected by their faces. To fold it up in the four dimensional space, one would have to fold along the 2-faces.

A way to give a formal definition of the unfolding of a polytope is the following. Define the adjacency graph of the polytope, that is a graph whose vertices are the facets of the polyhedron and an edge between two vertices is drawn if and only if the facets corresponding to the vertices are adjacent (that means, they have a common $(n-2)$-face, if n is the dimension of the polytope). For such a graph one can consider a spanning tree, that is a tree contained in the graph, which contains all the vertices of the graph. Every spanning tree is an unfolding of the polytope.

We describe two procedures for *projecting* an n-dimensional figure $F \in \mathbb{R}^n$ onto an hyperplane A. *Parallel projection*: given a line r not parellel to A, through each point of F draw a line parallel to r, to meet A in an "image" point. If r is orthogonal to A, then the projection is called *orthogonal*. *One-point perspective projection*: suppose a point $p \in \mathbb{R}^n$ (called projection center) is given such that each line passing through p and a point of F is not parallel to A. Then through each point of F draw a line that passes through p, to meet A in an "image" point.

For a polyhedron consider a graph whose vertices and edges represent respectively the vertices and the edges of the polyhedron, two vertices being adjacent exactly when the corresponding vertices of the polytope are endpoints of the same edge. It is not hard to see that every regular polyhedron has a planar graph, since one can project the polyhedron onto the sphere in which it is inscribed and then project the sphere onto the plane with a stereographic projection. Such a planar graph is called *Schlegel diagram* of the polyhedron. A Schlegel diagram of a polyhedron can also be obtained through a face-centered one-point perspective projection of the 1-skeleton of the polyhedron onto a hyperplane. Face-centered projection means that the line through the projection center and the center of a facet is orthogonal to the hyperplane of projection and the projection center is close enough to the face.

In an analogous way it is possible to define the Schlegel diagram of an n-polytope, and also in this case the Schlegel diagram can be obtained through a facet-centered one-point perspective projection.

The unit n-sphere is the set of points $S^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$. A generic *$n$-sphere* is the set of points $\{x \in \mathbb{R}^{n+1} : \|x - c\| = r\}$, where $c \in \mathbb{R}^{n+1}$ is the center and $r > 0$ is the radius of the sphere. We call hypersphere a 4-sphere.

4.2 Contents of the exhibition

The main subject of the exhibition is the hypercube. This polytope has been chosen for variuos reasons: it is one of the simplest regular polytopes, its properties can be compared with the analogous ones of the cube and finally it has some characteristics that allow to present nontrivial problems and results.

A little part of the exhibition is dedicated to the other five regular polytopes and to the hypersphere.

The exhibition is subdivided into seven parts: introduction, models, symmetry, slices, dice, flip a cube, regular polytopes. Some interactive animations and short movies complete the exhibition, presenting the themes developed during the visit. Our animations are endowed with written explanations, while the other ones are not, some of them being only recreational or intended to give suggestions.

The proposed visit does not touch all the seven parts, but only the introduction, the models, one among the central ones and at the end the section about polytopes. This choice is due to the special arrangement inside the "Festival della Scienza": there are many little exhibitions and the visitor does not have much time to dedicate to each one, usually not more than 30–45 minutes.

We describe in some detail each part of the exhibition. The section "flip a cube", which is the subject of our work, will be described in detail in section 4.3.

4.2.1 Introduction

Some posters narrate in a catchy way an imaginary dialogue between two friends, one of them has just come back from a tour in the 4D space. Some phenomena that happen in the 4D space are recalled in the story, without describing anything in detail. The posters are not intended to explain anything, but only to involve the visitor.

A large model representing a projection of a 120-cell (one of the regular 4D polytopes) completes this introductory section.

4.2.2 Models

This section is the central part of the exhibition. Its aim is to describe some 3D models of the hypercube, mainly by analogy with 2D representations of the cube. The representations presented with models and explications are the following:

- models obtained unfolding the cube and the hypercube, figure 4.1
- parallel projections, figure 4.2
- perspective projections (Schlegel diagram), figure 4.3.

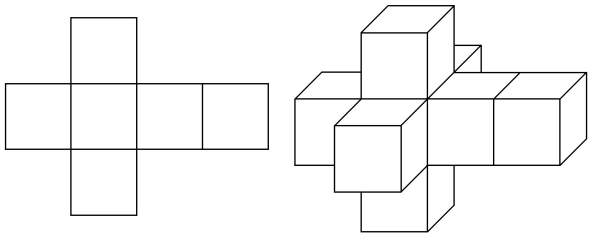

Figure 4.1: The unfoldings of a cube and a hypercube.

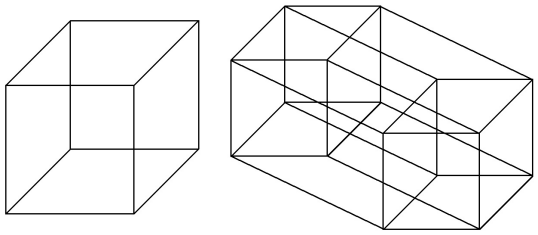

Figure 4.2: Parallel projections of a cube and a hypercube.

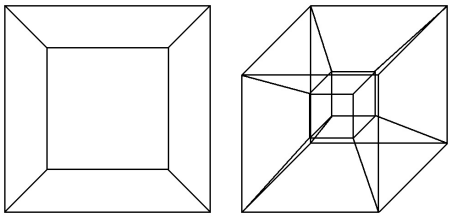

Figure 4.3: Schlegel diagram of a cube and a hypercube.

There is also a model of two polyhedral tori (figure 4.4): it represents two complementary solid tori of the hypercube and suggests with different colours which faces are to be identified. Each torus of the hypercube is made of four cubes in the 4D space, but our model deforms them to represent the linked tori in the 3D space.

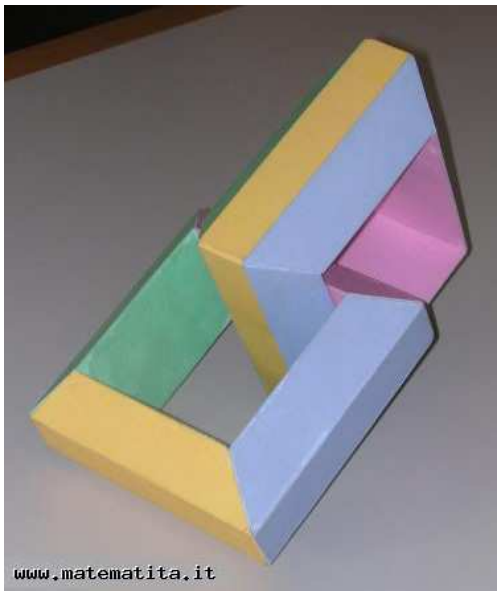

Figure 4.4: A model of the two tori of a hypercube (image by *matematita*).

Two problems are posed to help the visitor think about the different representations and their properties. To find the solution it can be useful to consider the different models at disposal.

The first problem asks to find, among some different constructions made of 8 cubes each, which ones can be obtained unfolding a hypercube. In other words, the visitor is invited to think about some necessary properties that the constructions of cubes must have to be folded into the hypersurface of a hypercube. The answer can be found comparing this problem with the analogous one for the cube. To complete this section, two posters show the 11 unfoldings of the cube among all possible plane constructions of six squares glued at some edges, and the 261 unfoldings of a hypercube.

The second problem invites the visitor to think about a path on the surface of the cube, that touches four faces forming a plane torus, i.e. the faces of the cube excluding two parallel ones. Then the analogous problem in the hypercube is considered. In the polytope the problem is more interesting, because the torus one can identify is solid, made of four cubes, and its complementary is again a solid torus made of four cubes, linked with the first one. In this way one can have an idea of the structure of two linked polyhedral tori represented in the model of figure 4.4. This concept is however difficult for the average visitor and it is not explained in the exhibition, but only mentioned and recalled by some models showing the tori. A particularly interested visitor can obtain more explanations from the guides.

4.2.3 Symmetry

This section investigates the rotation group of the cube, giving attention to the fact that it acts transitively on vertices, edges and faces. In other words, vertices, edges and faces are not distinguishable: in any way one fixes two of them, it is possible to take one into the other with rotations.

Thanks to some kaleidoscopes, the symmetry of the cube is highlighted and its rotation axes are pointed out. They can be seen also in figure 4.5: one axis of order 4 through the centre of two opposite faces, one of order 3 through two opposite vertices and one of order 2 through the centres of two opposite edges.

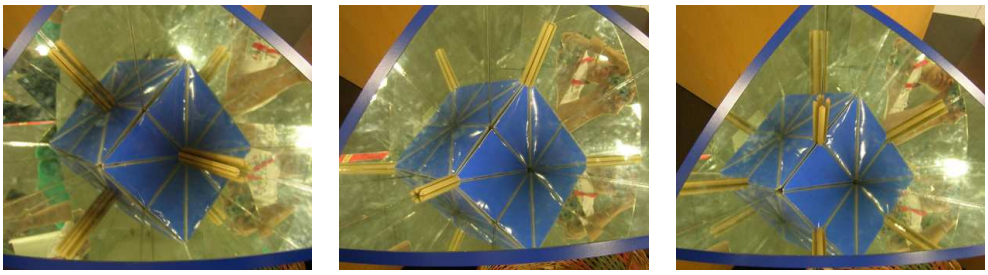

Figure 4.5: The rotation axes of the cube in a kaleidoscope.

The visitor should note that these axes are not easy to see in the two dimensional representations of the cube. Analogously, the three dimensional representations of the hypercube do not point out its symmetry.

The aim of this section is not to study the symmetry of the hypercube or to talk about symmetry planes of 4D objects, but only to convince the visitor that also the hypercube is "somehow" symmetric, even if that is not easy to note simply looking at its 3D representations.

4.2.4 Slices

Some models present different slices of the cube. The visitor is invited to use them to find out which polygons, among some given ones, can be obtained as slices of a cube. This problem is in general not trivial, but the polygons are chosen such that one needs only some evident properties to tell whether they can be slices of a cube.

This question is posed to prepare the visitor to an interactive animation, that shows the 3D slices of a hypercube along four hyperplanes.

4.2.5 Dice

How many ways are there to dispose the numbers on the faces of a cubic die?

The visitor has thirty different dice at disposal and is asked to tell whether they represent all the possible cases. One way to arrive to the solution is the following. Fix a face, giving it the same number for all dice and consider the opposite face: there can be five possible numbers on it. Now fix a lateral face and consider the opposite one, for which there are three possibilities. To dispose the numbers on the last two faces, there are other two possibilities. Then, there are 30 ways to dispose the numbers on the faces of a die.

Using the same argument, the visitor can be convinced that there are 210 different ways to dispose the numbers on the hyperfaces of a hypercube.

4.2.6 Flip a cube

This part is the subject of our work and is described in detail in the next section 4.3.

4.2.7 Polytopes

The visitor is helped to understand how many regular polyhedra there exist, by disposing regular polygons around a vertex and checking when they can be folded up in the 3D space. The visitor should understand that there are at most five regular polyhedra: one can dispose three equilateral triangles around a vertex (obtaining the tetrahedron), four equilateral triangles (octahedron), three squares (cube) or three pentagons (dodecahedron).

Analogously, the visitor can try and dispose regular polyhedra around an edge, without filling the whole space around that edge. Six possibilities will be found: three, four or five tetrahedra around each edge (that gives the hypertetrahedron, the hyperoctahedron and the 600-cell respectively), three cubes (hypercube), three octahedra (24-cell) or three dodecahedra (120-cell).

Some models suggest a way to obtain a polytope from one in a lower dimension: one considers the construction of the pyramid (to obtain a tetrahedron from a triangle and a hypertetrahedron from a tetrahedron), of the prism (to obtain a cube from a square and a hypercube from a cube) and of the bipyramid (to obtain an octahedron from a square and a hyperoctahedron from an octahedron). Here it is suggested that tetrahedron, cube and octahedron have a corresponding polytope in every dimension (and in fact they are the only regular polytopes that exist in every dimension, but this is not described).

4.3 Flip a cube

The phenomenon we want to explain is mathematically elementary: $SO(3)$ preserves the orientation of $\mathbb{R}^3$, but one can act with $SO(4)$ on a three dimensional object (that is, contained in $\mathbb{R}^3 \subseteq \mathbb{R}^4$) such that the image is again in $\mathbb{R}^3$, but has the opposite orientation. In other words, it is possible to realize a reflection of $\mathbb{R}^3$ as a rotation of $\mathbb{R}^4$.

The group of rotations $SO(3)$ of $\mathbb{R}^3$ is a subgroup of the group of rotations of $\mathbb{R}^4$ via the following mapping:

$$\begin{array}{ccc} SO(3) & \hookrightarrow & SO(4) \\ A & \mapsto & \left(\begin{array}{c|c} A & 0 \\ \hline 0 & 1 \end{array}\right) \end{array}$$

Also the set $O(3)^- = O(3) \setminus SO(3)$ of orthogonal 3×3 matrices with determinant -1 can be seen as a subset of $SO(4)$:

$$\begin{array}{ccccc} \psi: & O(3)^- & \hookrightarrow & SO(4) \\ & A & \mapsto & \left(\begin{array}{c|c} A & 0 \\ \hline 0 & -1 \end{array}\right) \end{array}$$

Then, since $SO(4)$ is connected, there is a rigid movement M_t which takes the identity matrix in $\psi(A)$, i.e. such that $M_0 = \mathrm{id}$ and $M_1 = \psi(A)$. In other words, a three dimensional object can be rotated in $\mathbb{R}^4$ and brought back into $\mathbb{R}^3$ such that is is flipped. It is then possible to flip a cube, but one can not notice when it is flipped, because of its symmetry, described also in a section of the exhibition. Hence we need to introduce an orientation on the cube to note when it is flipped. To do that, we use different colours, arranged clockwise or counterclockwise around a vertex.

We chose to "flip" a cube, because it can be moved in the hypersurface of a hypercube, and cubes and hypercubes are the principal subject of the whole exhibition, which the visitor should be familiar with, after visiting the first part of the exhibition. In order to explain the movement of the cube we use the analogy, treating first the case of flipping a square in the three dimensional space.

Let us consider a square first. To orientate a square it suffices to paint its edges with at least three different colours, to distinguish two edges which concur in a vertex. In the animations and in the models for the exhibition four colours were used. The reason is purely aesthetic and "psychological": most people who saw the models painted with three colours asked why two edges were painted with the same colour. That way of colouring seems to be perceived as "innatural" in some sense by many people.

The visitor can observe and use a model of a square coloured in this way. The orientation does not change if one simply rotates the square maintaining it adherent to the plane of the table. However, moving the square in the 3D space, it can be brought back onto the plane with the opposite orientation.

This movement can also be performed keeping the edges of the square always adherent to the surface of a cube, having the same edges' length as the square. The visitor is asked to try and find this movement, using the object shown in figure 4.6, a plexiglas cube with a metallic square slightly larger than the faces of the cube. The square remains adherent to the surface of the cube thanks to some magnets placed in the cube. Figure 4.7 gives an idea of the movement that filps the square.

Figure 4.6: The cube present in the exhibition.

Since a cube can be represented in two dimensions, also the movement of the square on the surface of the cube can be drawn in two dimensions. This step is necessary in order to understand the corresponding movement for a cube in the hypercube. In fact, for a square on a cube we can see both the 2D representation and the 3D object, while when we consider the hypercube, we can at best see one of its 3D representations and one can only try and imagine what happens in the 4D space.

The visitor can then use the computers where some interactive animations help him to be aware of the phenomenon. The movement of the square is represented in two 2-dimensional models of the cube (its unfolding and its Schlegel diagram: see figure 4.8) and the movement of the cube in the hypercube is represented in the corresponding models of the hypercube (figure 4.13). We describe these animations in section 4.3.2 and 4.3.3.

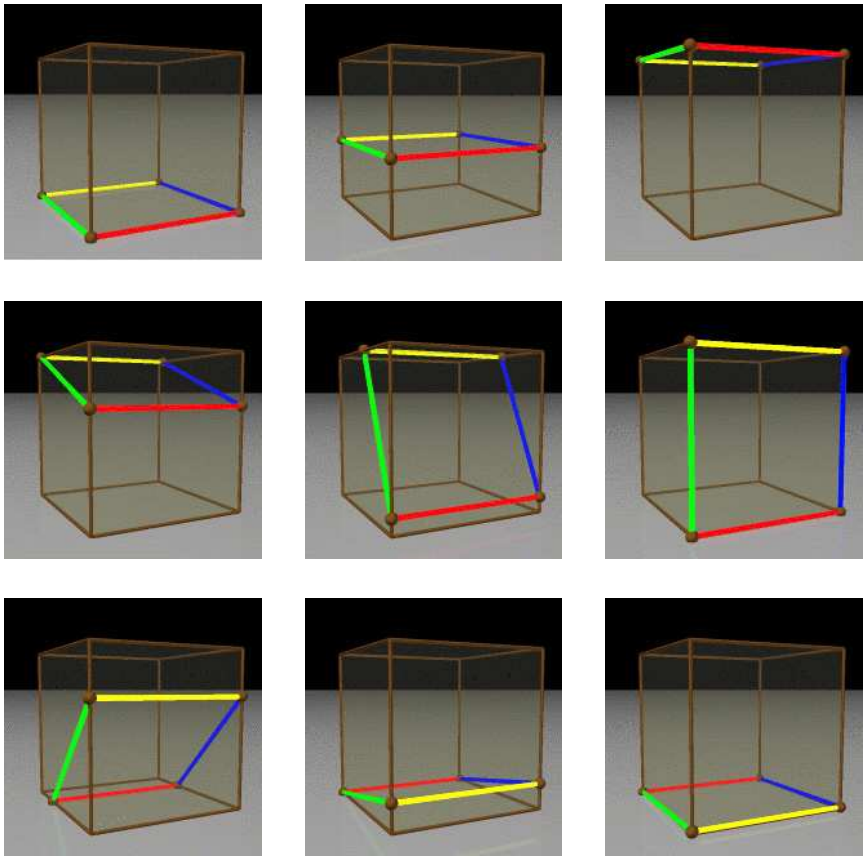

Figure 4.7: How to flip a square on the surface of a cube.

4.3.1 Animations and images

To develop the animations and to create animated images we used free software:

- Java™Platform, Standard Edition 6 Development Kit (JDK™)[2], a development environment for building applications, applets, and components using the Java programming language;
- JoGL[3] (Java OpenGL), a wrapper library to use OpenGL in the Java programming language;
- NetBeans IDE 5.5[4], a free, open-source Integrated Development Environment for software developers;

[2] `http://java.sun.com/`
[3] `https://jogl.dev.java.net/`
[4] `http://www.netbeans.org/`

- POV-Ray[5] (Persistence of Vision Raytracer), a free tool (with source code available) for creating three-dimensional graphics.

Nine interactive animations were developed in the Java programming language. Four of them describe how it is possible both to flip a square on a cube and a cube on a hypercube.

The other ones are about related topics, and are described in the next section 4.4. Two animations illustrate some two dimensional representations of a cube and some three dimensional representations of a hypercube. Two animations show different projections of a hypercube. The last one describes another phenomenon: how to reverse a sphere in a hypersphere.

Some common characteristics of the interactive animations are listed here.

- There is a control panel under the drawing. It allows the user to interact with the animation in an immediate way.
- One or more sliders control the parameters (such as angles of rotation or frame of the animation) that manage the main actions.
- A button lets the animation play and stop (not available in some applets).
- Other buttons control specific parameters for the applet, for example to choose the model to display.
- The left mouse button controls a trackball to rotate the object.
- The right mouse button controls the zoom, which changes the size of the figure.

Some animated images were realized using *POV-Ray* to support the description of the interactive animations. Some images show the analogous phenomenon as an animation, but in a lower dimension. Some images show exactly the same as an animation, and they are useful as short films to be placed in the exhibition. They are intended not necessarily to explain something, but also to give some suggestions and to make the exhibition more interesting. The advantages of these images versus the interactive animations are the better quality of the 3D graphic and the lower use of cpu.

The applets are organized into web pages. Each page contains both the instructions to interact with the applet and a brief informal explanation of what is going on from a mathematical point of view. The latter text contains also links to the other pages.

[5] `http://www.povray.org/`

The web pages can be found in the CD–ROM included to this thesis, where also the source code and the stand alone applications are available.

4.3.2 Flip a square

The animations in figure 4.8 show the square moving on the surface of a cube till it is flipped. The cube is represented by one of its unfoldings in the first animation and by the Schlegel diagram in the second, i.e. a central perspective projection where the centre of the projection is located near the centre of the face that is parallel to the projection plane and far from it. Central projection means that the projection plane is parallel to two faces of the cube. In other words, the Schlegel diagram is a graph whose vertices represent the vertices of the cube and the edges, that represent the edges of the cube, meet only in the vertices.

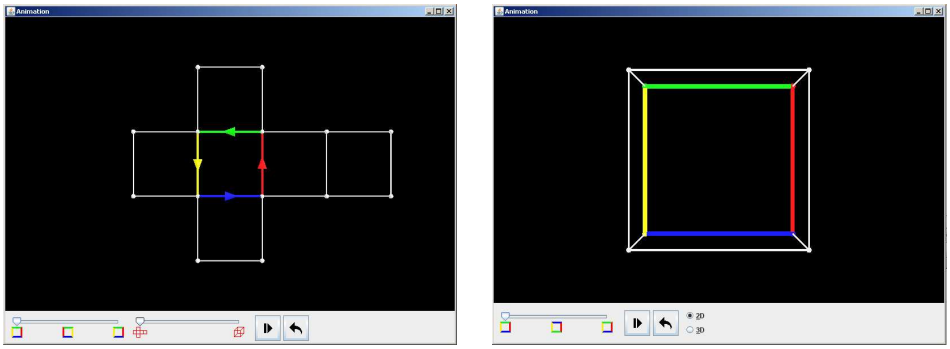

Figure 4.8: Flipping a square on a cube.

In both animations it is possible to see the object as two dimensional or as a representation which gives the idea of 3D.

We report the text written for the web page containing the first animation.

> Moving a square on a plane, it is not possible to flip it. Paint its edges yellow, blue, red and green, disposing the four colours clockwise along the perimeter. Any movement of the square on the plane lets the colours be disposed in the same way.
>
> However, if the square is free to move in the 3D space, it is easy to flip it and bring it back onto the plane with the colours disposed counterclockwise.
> It is also possible to flip the square without detaching its edges from the surface of a cube.
> If the cube is represented by one of its unfolding, it is then possible

to follow the movement of the square on a plane, even if it takes place in the 3D space.

The square seems to be broken and in some moments the edges are painted twice, but folding up the cube one can verify that the segments always form a square.

In an analogous way, it is not possible to flip a cube in the usual space, but that becomes possible in the 4D space and one can follow the movement on an unfolded hypercube.

It is possible to see the movement of the square also in another plane representation of the cube, its Schlegel diagram.

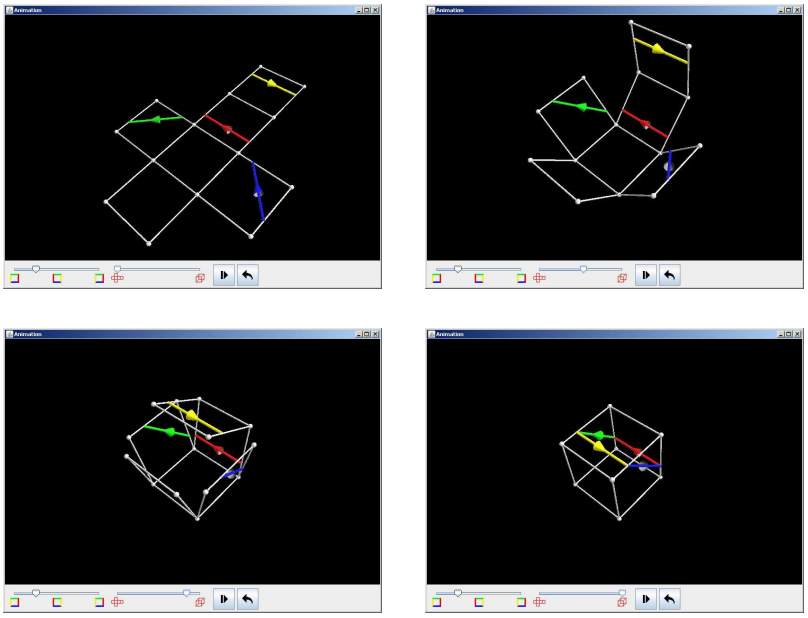

Figure 4.9: Folding up the cube to see the square.

The user should note that in the two dimensional representations some properties of the 3D object are not preserved.

In particular, in the unfolded cube, the edges of the square do not seem to be always adjacent and the square seems to "open". In fact the cube can be seen as quotient of its unfoldings, where some of the edges must be identified following the relations which permit to fold up the model in the 3D space. Another property of this representation that can be noted by the user is that the edges of the square are drawn twice when they coincide with the cube edges that must be identified. The user can comprehend that these

"errors" are due to the representation we choose for the cube: the model can not preserve all the properties of the cube because it is contained in a space with lower dimension as the cube itself. In this sense it is important that the user can "close" the unfolding to obtain the cube and verify that the figure he sees is always a square.

Figures 4.9 and 4.10 show two examples where the unfolded cube is folded up to obtain the 3D object: in this manner the user can verify that the figure drawn is always a square, even if it does not appear so in the 2D representation.

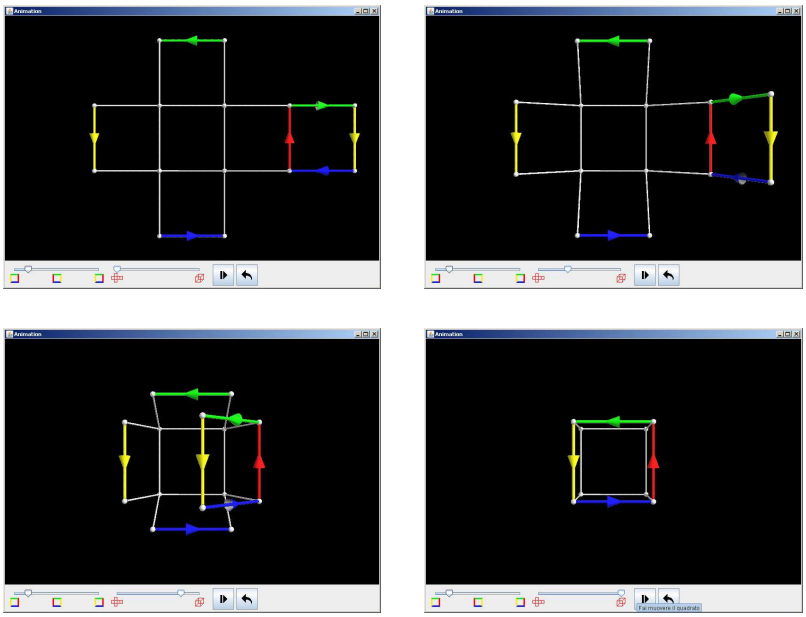

Figure 4.10: Folding up the cube to see the square.

In the second animation, the default view mode is 2D, that means the projection of the cube is fixed (Schlegel diagram), but one can switch to a 3D view, activating the trackball to make possible to change the projection centre and see different projections. The shading helps interpret the objects as three dimensional.

In the Schlegel diagram there are representation "errors" as well: the square becomes a trapezium and even a line segment when the point of view is complanar with the square. Moreover, the edges length varies during the movement. Examples of these drawings can be see in figures 4.11 and 4.12, where the 2D view is shown together with the 3D view from the same point of view and from another point, to make easier to comprehend how the 3D figure looks like.

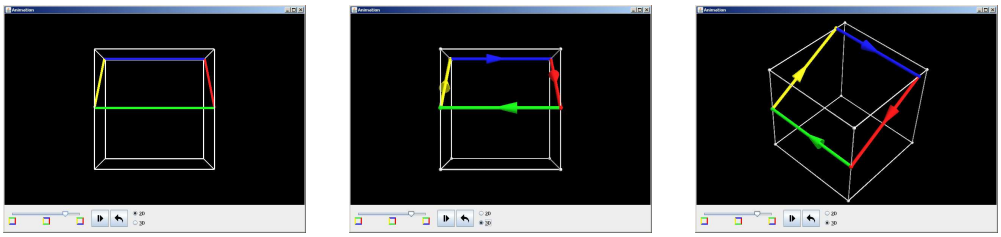

Figure 4.11: A square seems a trapezium in the Schlegel diagram.

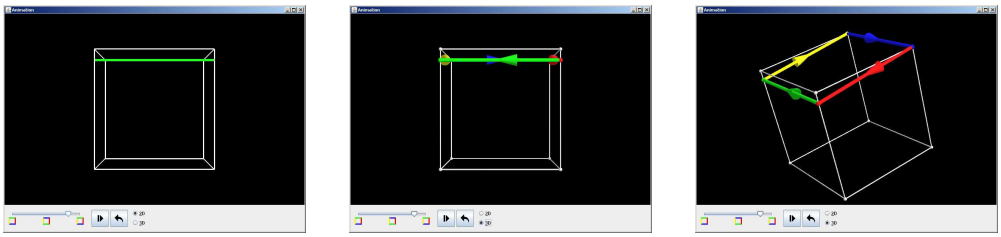

Figure 4.12: A square seem a line segment in the Schlegel diagram.

The text for the web page is almost the same as for the previous animation, but the paragraph on the characteristics of the 2D model is substituted by the following.

> The square seems to become larger and smaller, deform into a trapezium and degenerate into a line segment, but looking at the cube from different positions, it is possible to verify that the quadrilateral is in fact always a square.

The visitors are now prepared to see what happens in a higher dimension: they have just seen that a 2D model of a 3D object might cause some representation "errors", and can now accept that the same holds for 3D models of 4D objects.

4.3.3 Flip a cube

The animations in figure 4.13 show how a cube can be flipped moving in the hypercube. The usual two representations of the hypercube are used: the first animation considers some different unfoldings, while the second one reproduces a Schlegel diagram of the hypercube.

In the animation with the unfolded hypercube, we chose to let the user change the unfolding, to see how some of the 2-faces are identified, in order to verify that the cube faces are always adjacent even if the cube seems to be "open" and sometimes some faces are drawn twice. These are exactly the

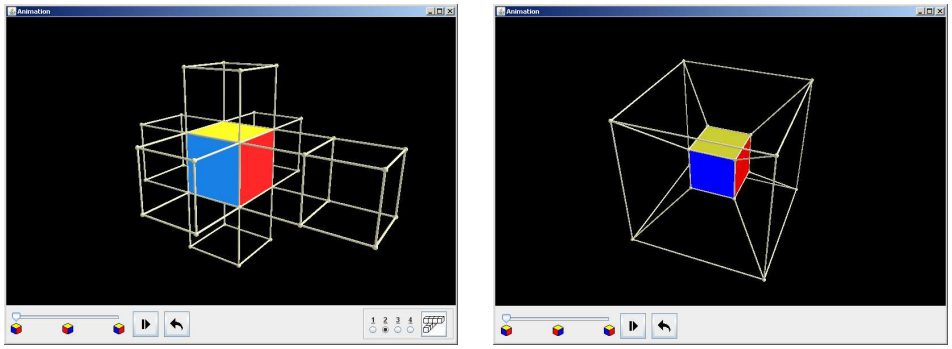

Figure 4.13: Flipping a cube on a hypercube.

same "errors" as for the square in the cube and can be seen in figure 4.15. Some of the infoldings of the hypercube are shown in figure 4.14, together with the intermediate movements that take an unfolding in the other.

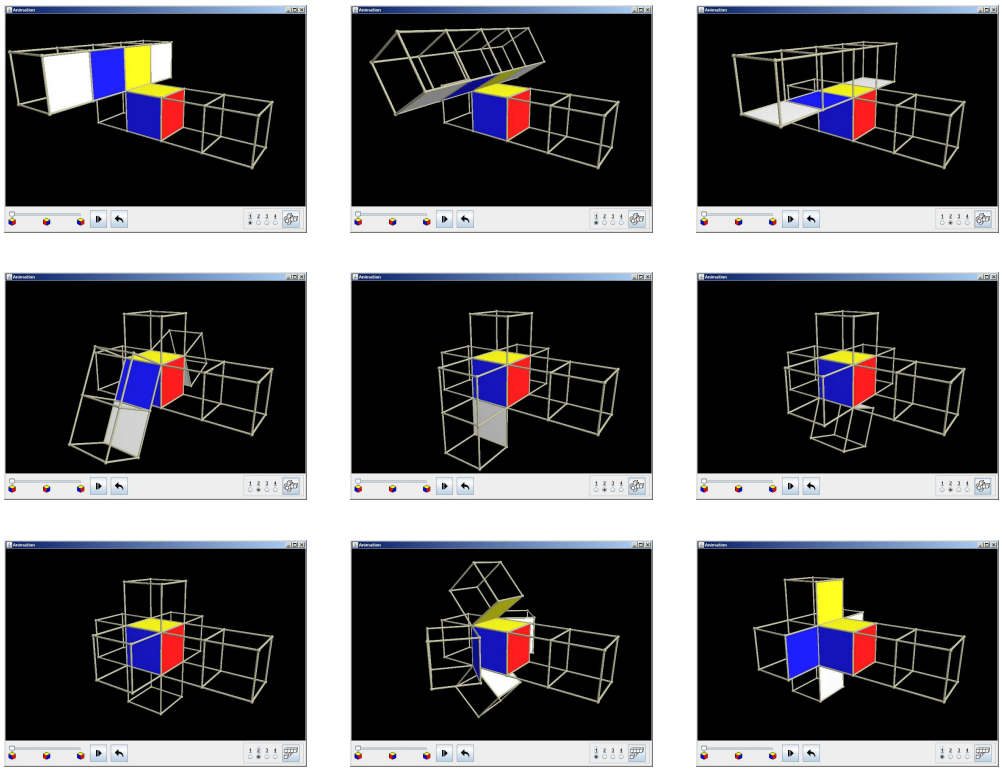

Figure 4.14: Changing the unfolding of the hypercube.

The text written for the web page containing this animation is the following.

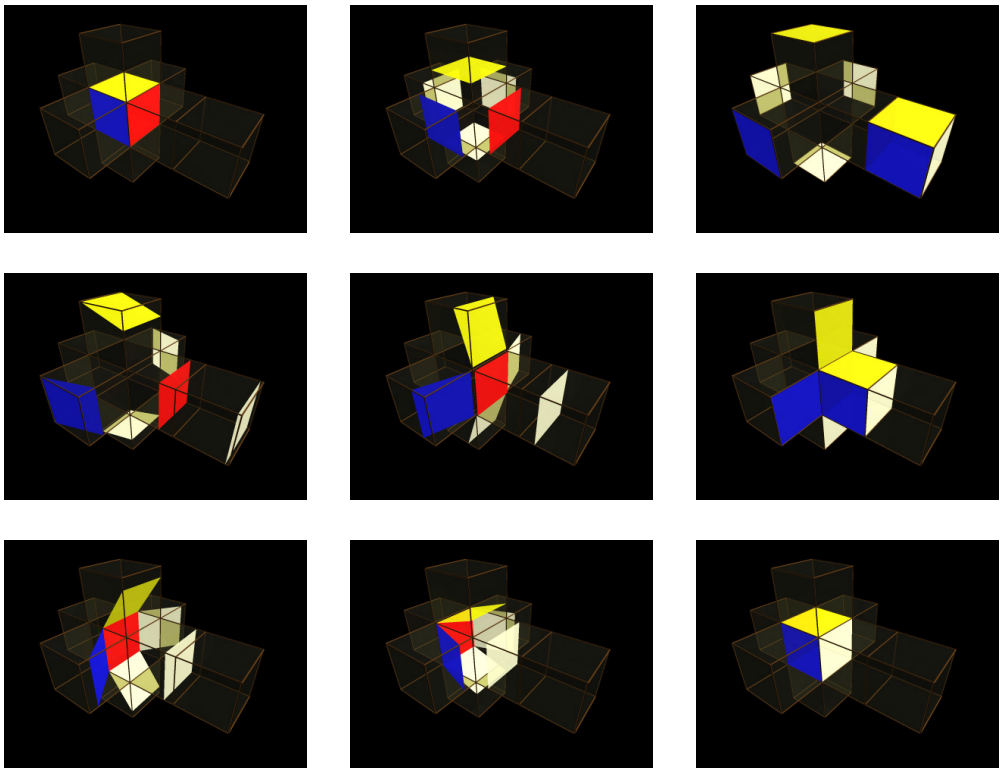

Figure 4.15: How to flip a cube on an (unfolded) hypercube.

Moving a cube in the usual 3D space, it is not possible to flip it. Paint three faces - concurrent in a vertex - yellow, blue and red, disposing the colours counterclockwise around the chosen vertex. Any movement of the cube in the space lets the colours be disposed in the same way.

As for a square, if we move the cube in the four dimensional space, we can flip it and bring it back to our space with the colours disposed in the opposite way.
It is possible to flip the cube without detaching its faces from the hypersurface of a hypercube.
If the hypercube is unfolded, it is then possible to follow the movement of the cube in our space, even if it takes place in the 4D space.

The cube seems to be broken and in some moments the faces are painted twice, but changing the unfolding of the hypercube, one can verify that the six faces are always adjacent.

It is possible to see the movement of the cube also in another 3D

representation of the hypercube, its Schlegel diagram.

To understand the relation between the two models of the hypercube, an animation shows how to fold up and unfold the hypercube.

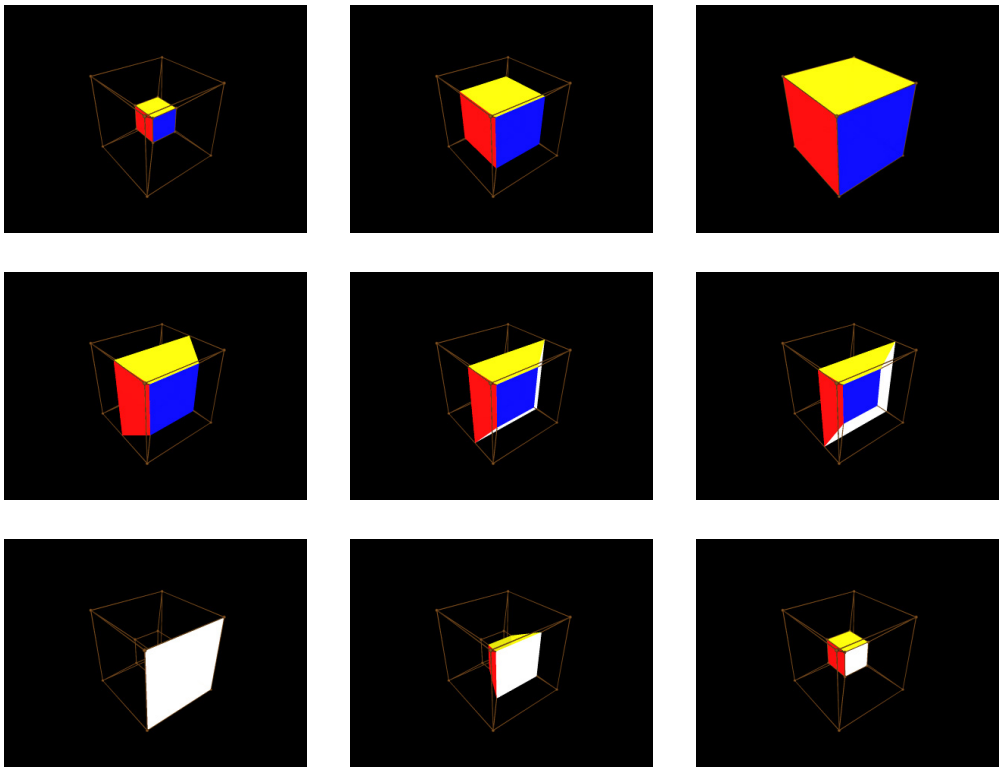

Figure 4.16: How to flip a cube on a hypercube (Schegel diagram).

The text written for the web page containing the second animation is almost the same but the paragraph about the properties of the 3D model is substituted by the following, which describes the "errors" in the representation of the cube that can be seen in the Schlegel diagram. They can be noted also looking at figure 4.16.

> The cube seems to become larger and smaller, deform into a truncated pyramid and degenerate into a square. That happens because we are representing a 4D object in the 3D space: as some of the cubes in the hypercube, also the cube is deformed by the projection from the 4D to the 3D space.

The same movement of the cube in an unfolded hypercube and in the Schlegel diagram is the subject of two animated images, some frames of which can be seen in figure 4.15 and 4.16.

4.4 Other animations

Other interactive animations were developed to illustrate topics related to the previous one.

4.4.1 Folding up a (hyper)cube

The animations in figure 4.17 should help the visitor to understand how the two representations of the hypercube (the unfolding and the Schlegel diagram) are linked: we can apply the same projection which gives the Schlegel diagram also to the hypercube while we fold and unfold it.

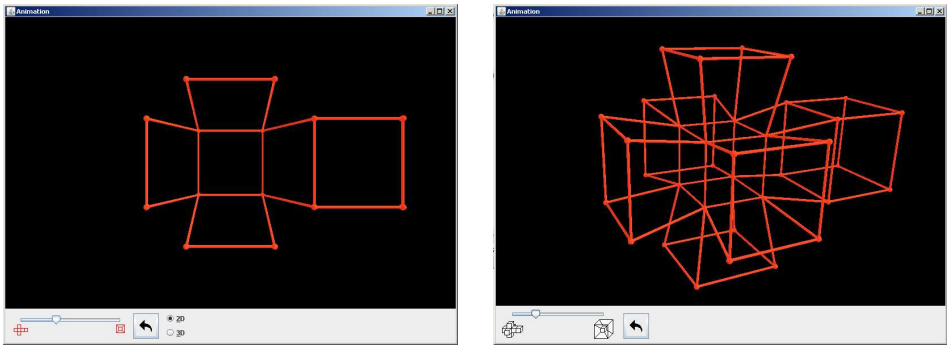

Figure 4.17: Folding a cube and a hypercube.

The first animation is a two dimensional projection of an unfolded cube that folds up in the 3D space to construct the cube. In the default view the unfolding lies in a plane that is parallel to the projection plane, but the user can rotate the object.

The same construction is the subject of an animated image, figure 4.18 shows some of its frames.

The text written for the web page containing this applet is the following.

> There are different ways to represent a cube in the plane, for example unfolding it (there are 11 different plane unfoldings) or projecting it to obtain the Schlegel diagram.
> If we observe the "most classical" 2D unfolding of a cube, the latin cross, that folds up in the 3D space, we obtain a two dimensional representation of the passage from a model to the other.
>
> First the four faces adjacent to the basis are folded up till they reach a vertical position, then the lid is shut.

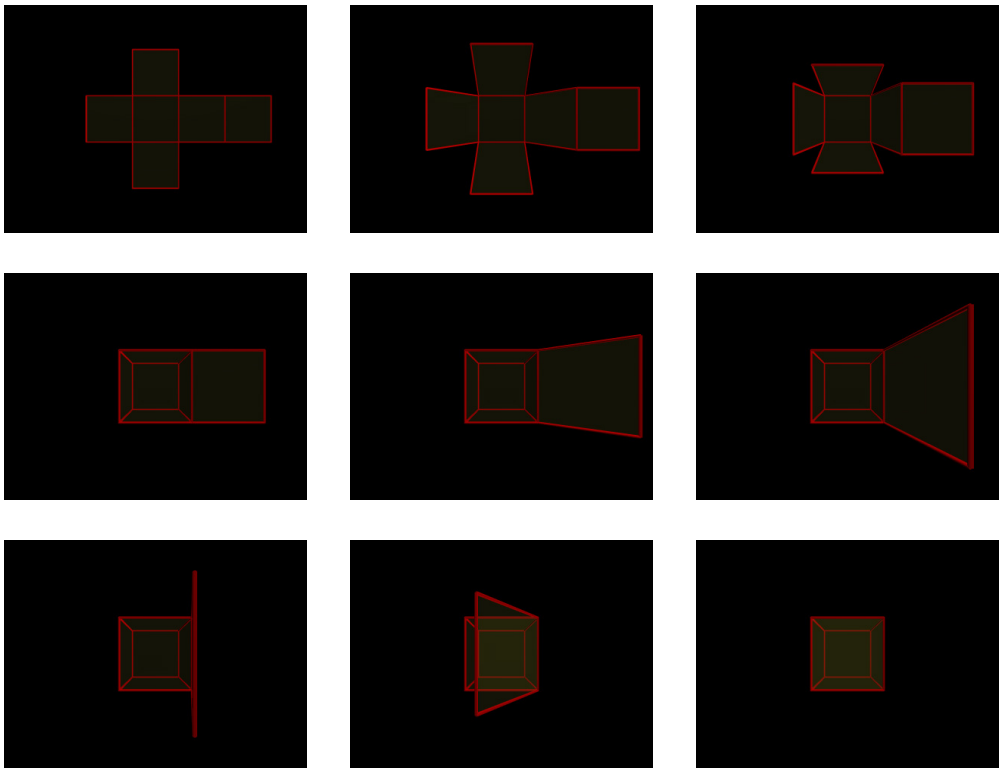

Figure 4.18: How to fold a cube in the 3D space.

> In an analogous way, one can represent the 3D unfolding of a hypercube folding up in the 4D space, to obtain the Schlegel diagram of the hypercube.

The second animation is the analogon of the first one, in a higher dimension. It shows how a three dimensional unfolding of a hypercube can be folded up in the four dimensional space to obtain the 4D object. The 4D object is projected into the 3D space.

In this animation two projections are applied to the hypercube: the first one maps the 4D object into the 3D the space, and the second one projects the 3D object onto a 2D representation on the monitor.

The user can change the centre of the second projection, that is rotate the 3D model, but the first projection is fixed, the centre being close to the centre of a hyperface.

The text for the corresponding web page is the following:

> There are different ways to represent a hypercube in the 3D space, for example unfolding it (there are 261 different solid unfoldings) or

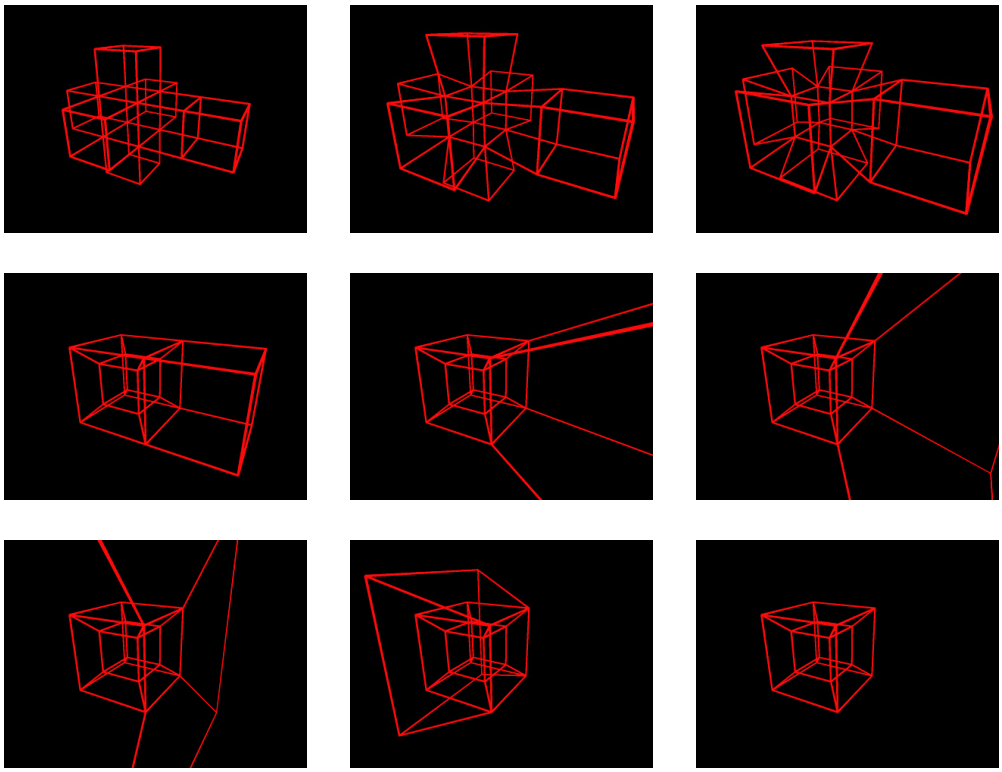

Figure 4.19: How to fold a hypercube in the 4D space.

projecting it to obtain the Schlegel diagram.
As for a cube, also the hypercube can be "constructed" folding up in the 4D space one of its three dimensional unfoldings.

Through a projection, we can represent this construction in the 3D space. First the six hyperfaces adjacent to the basis are folded up till they reach a "vertical" position, then the lid is shut.

An animated image has the same subject, some frames of which are in figure 4.19.

4.4.2 Projections

The two animations in figure 4.20 show different projections of a hypercube.

The first one allows the user to act on a slider, controlling the projection centre: it is at first near the hypercube (giving a perspective projection), but can be made far away from the object, since one obtains a parallel projection.

To understand what happens with the centre of projection, the user can look at an image that describes the projection of a cube onto a plane. Some

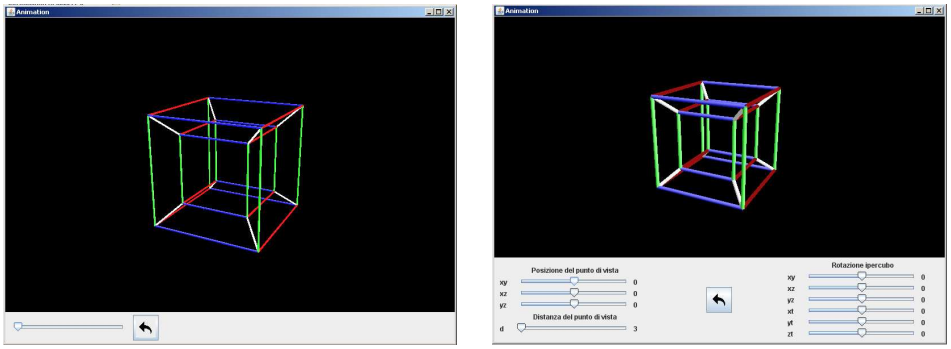

Figure 4.20: Projections of a hypercube.

frames of this image are in figure 4.21. Here the projection centre and the rays through the cube vertices are drawn, and define the cube shadow on the projection plane.

The web page contains the animated image of figure 4.21 and the following text.

> To obtain a two dimensional representation of the cube, it is sufficient to project it onto a plane. We can choose the plane to be parallel to a face of the cube (one-point perspective). The image obtained through the projection changes according to the projection centre. If this is located near the centre of a face, we obtain the Schlegel diagram of the cube, where
>
> - the face closest to the plane seems to be the largerst one and to contain all the other ones;
> - parallel edges that are also parallel to the plane are represented by parallel segments;
> - parallel edges that are not parallel to the plane are represented by segments that converge to a point.
>
> If we move the projection centre far away from the cube, till it reaches an "infinite distance", the projection rays become parallel and we obtain an axonometric projection, where
>
> - the faces parallel to the plane are projected onto equal squares;
> - parallel edges are represented by parallel segments.
>
> In an analogous way we can represent the hypercube by projecting it into the usual 3D space. We obtain different representations according to the projection centre. When this is close to the centre

of a hyperface, we obtain the Schlegel diagram, while if it is infinitely far from the hypercube, we have a parallel-edges representation.

To discover more about this topic, you can look at other projections, obtained by rotating the hypercube in the 4D space or moving the projection centre in different ways.

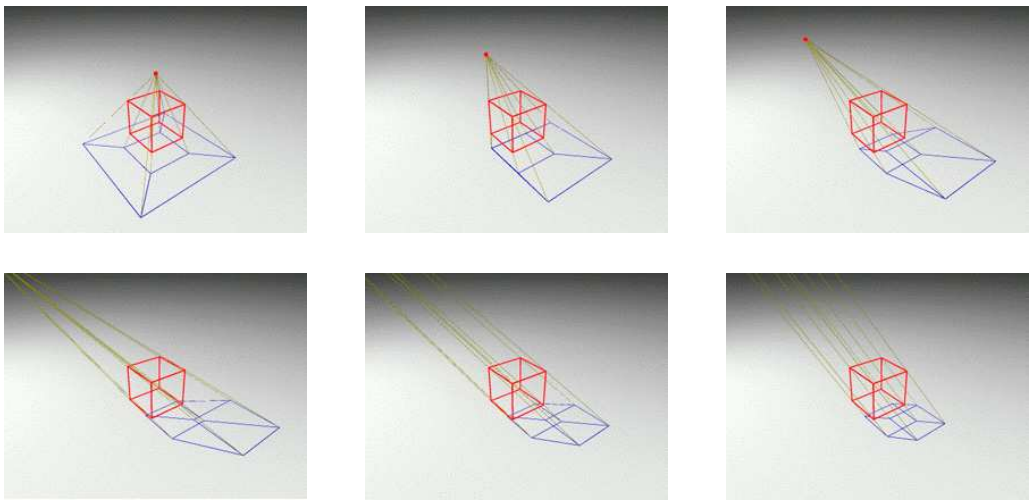

Figure 4.21: Projections of a cube.

Some frames of this animation can be seen in figure 4.22: the first frame shows the Schlegel diagram, a central projection (or one-point projection), while the last one shows an axonometric projection.

The edges of the hypercube are parallel to the coordinate axis in the 4D space. The edges that are parallel to each other are drawn with the same colour. Since the space of projection is parallel to three of the coordinate axes, the edges along these three directions are parallel also in the central projection, while the edges that are parallel to the fourth axis appear as convergent to a point. The two cubes parallel to the 3D space of projection are not equal in the projection: the one closest to the space of projection is represented by a smaller cube because it is far from the projection centre. In the axonometric projection the projection rays are parallel to each other. Every group of edges, which are parallel to an axis, is projected onto parallel segments and the two cubes parallel to the 3D space of projection are projected onto equal cubes.

The meaning of the coloration is explained in the text that accompanies the next animation. There we also mention what happens to parallel edges, but we are not expecting that every visitor comprehends this concept. In

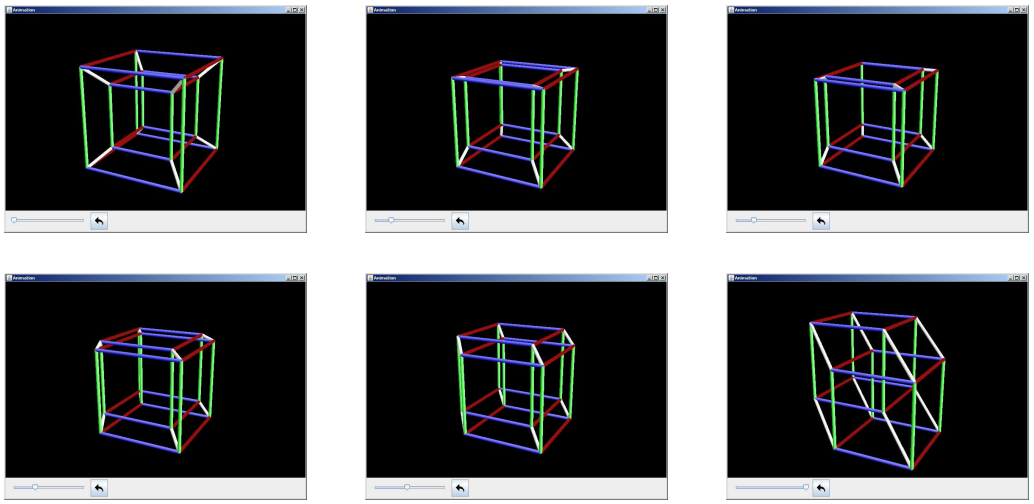

Figure 4.22: Projections of a hypercube.

fact the second animation concerning the hypercube projections is probably more difficult to use and understand for a visitor since more parameters can be controlled: the position of the projection centre and the position of the hypercube with respect to the 3D projection space.

We recall a simple fact about rotations in order to describe how the animation is constructed: we claim that the group $SO(4)$ of rotations in $\mathbb{R}^4$ is generated by the rotations about the six distinct planes each of whose is determined by two coordinate axes.
In fact, suppose R is an element of $SO(4)$. Note that if R fixes the first coordinate axis, $Re_1 = e_1$, then proving our claim is equivalent to say that $SO(3)$ is generated by the rotations about the three coordinate axes. This is a well known fact. Let $Re_1 = v \neq e_1$. To prove our claim, it suffices to find subsequent rotations about the coordinate planes that map $v = (v_1, v_2, v_3, v_4)$ to e_1.

Note that one can map (a, b) to $(*, 0)$, defining

$$(x, y) \mapsto \begin{pmatrix} \frac{a}{\sqrt{a^2+b^2}} & \frac{b}{\sqrt{a^2+b^2}} \\ -\frac{b}{\sqrt{a^2+b^2}} & \frac{a}{\sqrt{a^2+b^2}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

This is a rotation, because the matrix has determinant 1. It follows that we can map v to e_1 through subsequent rotations, in the following way. There is a rotation $R^{\theta_1}_{e_2,e_3}$ such that $R^{\theta_1}_{e_2,e_3} v = (v_1', v_2, v_3, 0)$. A rotation $R^{\theta_2}_{e_2,e_4}$ maps $(v_1', v_2, v_3, 0)$ to $(v_1'', v_2, 0, 0)$. Finally a rotation $R^{\theta_3}_{e_3,e_4}$ maps $(v_1'', v_2, 0, 0)$ to

$(w_1, 0, 0, 0)$. Since the rotations preserve the norm, we have $w_1 = \pm 1$. If $w_1 = -1$, take $R^{\theta_3+\pi}_{e_3,e_4}$ instead of $R^{\theta_3}_{e_3,e_4}$.
Then $R^{\theta_3}_{e_3,e_4} R^{\theta_2}_{e_2,e_4} R^{\theta_1}_{e_2,e_3} v = e_1$ and the claim is proved.

From this fact follows that we can determine how the hypercube is rotated with respect to the 3D space fixing six angles. These angles are controlled by six sliders.

There is also the trackball, which rotates the 3D projection of the hypercube. It gives the idea of threedimensionality to the projection. If the projection centre is on the fourth axis then rotating the projection is the same as rotating the hypercube about the planes that contain the fourth axis. But when the projection centre is in a general position, this does not hold anymore.

Four more sliders control the position of the 4D projection centre, fixing three angles and the distance from the origin. The three angles define the line through the origin and the centre of projection, and the distance defines uniquely the position of the centre.

The text for the web page is the following.

> The projection of a hypercube in the three dimensional space depends on the position of the projection centre, but also on the position of the hypercube with respect to the 3D projection space.
>
> In the animation at first two hyperfaces are parallel to the 3D space, that has coordinates x, y and z. The edges are parallel to the coordinate axes and are drawn red, green, blue (those parallel to the axis x, y, z respectively) and white (those parallel to the fourh axis).
>
> The projection centre is close to the centre of a hyperface parallel to the 3D space. In this way we get the Schlegel diagram. The red edges are projected onto parallel segments, and the same holds for the green ones and the blue ones, while the edges parallel to the fourth axis are projected onto segment converging to a point.
>
> In the 4D space an object can rotate about a plane. If we rotate the hypercube about the xy plane, acting on the top right slider, the red segments remain parallel, and also the green ones, since the corresponding edges are still parallel to the 3D space after the rotation. On the other hand, the blue edges, that are no more parallel to the 3D space, are now projected onto segments converging to a point.
>
> When the projection centre is at infinite distance, we get axonometric projections and the edges with the same colour are in any case

projected onto parallel segments.

4.4.3 Hypersphere

The last animation developed for the exhibition is about a different object: the hypersphere. We want to show how a sphere can be reversed in the hypersphere: in contrast to the movement of the cube, here we use topological deformations, not only rigid movements. We want to make an inversion: bring the sphere into itself, in such a way that the inside is now outside and viceversa. Figure 4.23 shows a step in the movement of the sphere.

If we define the hypersphere S^3 as a subset of $\mathbb{C}^2$,

$$S^3 = \{(x, y) \in \mathbb{C}^2 \mid |x|^2 + |y|^2 = 1\}$$

then the set

$$S^3 \cap \left\{(x, y) \in \mathbb{C}^2 \mid |x|^2 \leq \frac{1}{2}\right\}$$

defines a solid torus. The torus $S^3 \cap \{|x|^2 \geq \frac{1}{2}\}$ defines a second torus, which meets the first in the common boundary. Hence we can represent the hypersphere by two linked solid tori.

The meridians of a torus must be identified with the parallels of the other and viceversa. The identification is helped by some meridians and parallels drawn with two different colours.

We chose this representation with two tori because it is analogous to one of the representations of the hypercube presented in the exhibition (figure 4.4).

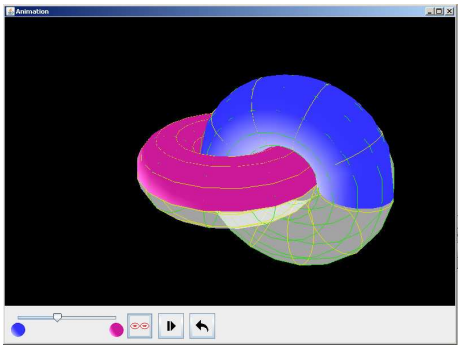

Figure 4.23: A (deformed) sphere inside the solid tori of the hypersphere.

The inside surface of the sphere is painted with a different colour as the outside.

We describe figure 4.24, which reports some frames of the animation. At the beginning the sphere is inside a torus, then is becomes larger till it touches the surface of the torus in one meridian. So it touches also the surface of the second torus, along a parallel.

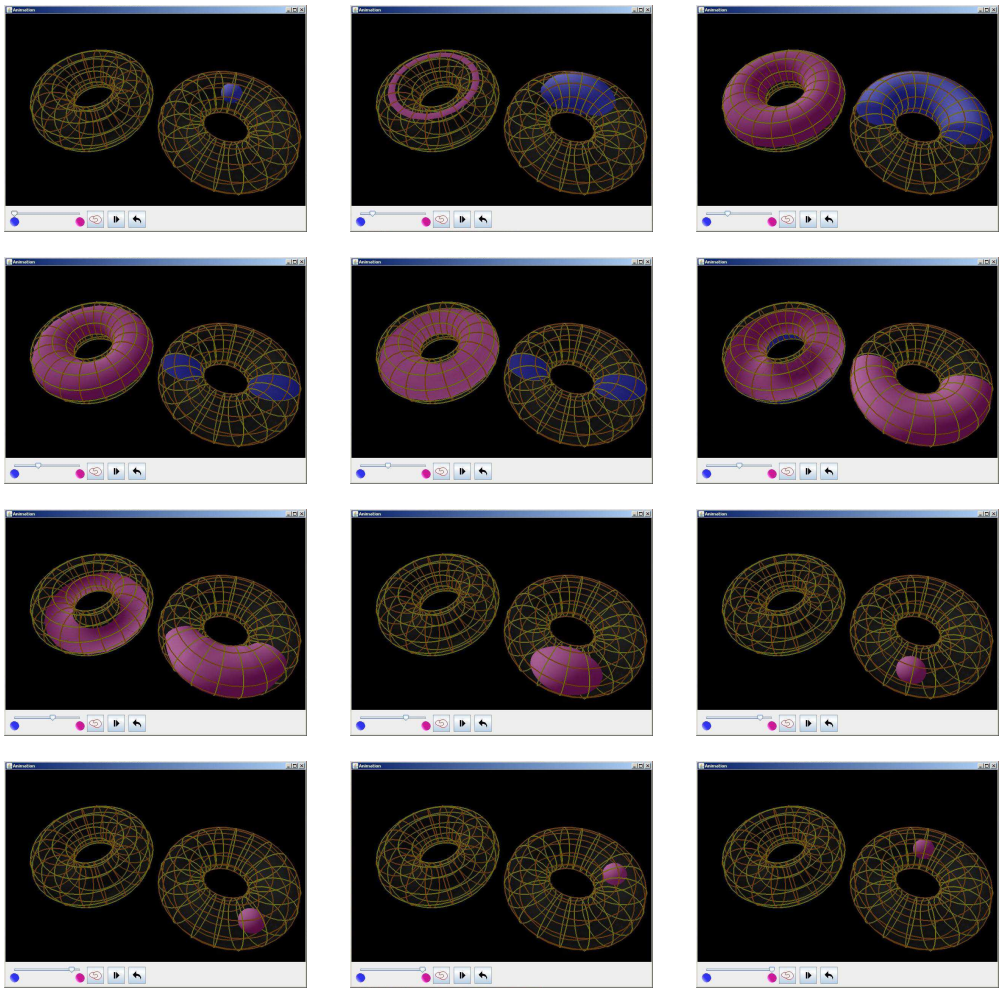

Figure 4.24: The sphere deformation in the hypersphere.

Then we inflate the sphere, that remains in the first torus till it becomes half the torus, delimited by two disks whose boundaries are two opposite meridians. Now we push the sphere into the second torus, till it becomes maximal, i.e. has the same radius as the hypersphere. In this moment the sphere "cuts" each torus into two symmetric parts. One torus is cut along two opposite meridians and the other along two opposite parallels.

We deflate the sphere in the first torus, exactly in the same way as we inflated it. The sphere is then entirely contained in the first torus and can

be moved inside it till it reaches the same position it had at the beginning. Now the inside of the sphere has become outside and viceversa: this can be noted observing the colour of the surface.

This is exactly the same as "inflating" a circumference on the surface of a sphere till it becomes maximal and then "deflating" it in the other hemisphere. The interior has now taken the place of the exterior.

Colouring the sphere is nothing else as choosing a vector field that is normal or transverse to the surface of the sphere. We chose a "complicated", not smooth, movement for the sphere, because we wanted it to have the surface on the surface of the tori for the longest time possible, in order to visualize it in our representation of the hypersphere. However, the phenomenon we are showing here can be obtained also with a diffeotopy (that is, a differentiable ambient isotopy). We describe a diffeotopy $f : S^3 \times I \to S^3$ that takes the 2-sphere $S \subseteq S^3$ into itself, exchanging its interior with the exterior. Note that such a diffeotopy exists: one can inflate the sphere maintaining it parallel to itself till it becomes maximal, then deflate it in the other hemisphere of the hypersphere till it has the same radius as at the beginning and finally apply a rotation of π about the plane parallel to the sphere and passing through the centre of the hypersphere. Then $f_t : S^3 \to S^3$ is a diffeomorphism for each $t \in I$ such that $f_0 = \mathrm{id}$, $f_1(S) = S$ and $f_{1|S} : S \to S$ reverses the orientation. Furhermore, $f_t(S)$ is a smooth sphere for every $t \in I$ (this is not true for the movement chosen in the animation).

We claim that $f_t : S^3 \to S^3$ always maintains the orientation. Suppose not, then there is a $t' \in I$ such that $df_{t'}(x) = 0$ for some x, which is a contradiction to the fact that f_t is a diffeomorphism for every $t \in I$.

Choose a vector field n that is normal to the surface of the sphere at $t = 0$. Then $df_t \cdot n$ is a transverse field to S in every moment, that means the coloration is defined in every moment by the normal component. Furthermore, $df_1 \cdot n = -kn$, where k is a positive function. In other words, the interior surface of the sphere has become the exterior and viceversa.

The text written to explain the animation is probably very difficult for the visitor.

> To describe the hypersphere, the 4D analogon to the sphere, we can think of two solid tori that are glued together in such a way that the meridians on the surface of one torus coincide with the parallels on the surface of the other one, and viceversa. In the animation meridians and parallels are painted with different colours to emphasize the identification.
>
> A sphere, painted blue outside and pink inside, is in the hypersurface

> of the hypersphere, inside one torus. We inflate the sphere till it touches the surface of the torus in a meridian. Since the surfaces of the tori are identified, the sphere appears also on the other torus, along a parallel.
>
> We inflate the sphere further and the part on the surface of the tori passes entirely inside the second torus, till the sphere is represented by two disks (in the first torus) which are adjacent to an annulus (inside the second torus).
>
> Now we can deflate the sphere symmetrically and it returns inside the torus where it was at the beginning, but painted pink outside. In other words, it is reversed: the inside surface is now outside and viceversa.

This text is surely the most difficult for the visitor. The main reason is that the exhibition does not present the hypersphere and in this page only a very short explanation about the representation of the hypersphere is given.

4.5 Future work

For the future we are planning to explain the construction of the hypersphere as two linked solid tori in analogy with the polyhedral solid tori in the hypercube (see figure 4.4). We also plan to talk about the (topological) "similarity" between cube and sphere and translate this in a higher dimension to give a better idea of a hypersphere.

Another idea is to present knots and links in three dimensions and to show that they can be untied in four dimensions. It would be also interesting to investigate how a sphere can be knotted in four dimensions. However this is probably too difficult to be presented in an exposition and explained to general public.

Acknowledgments

I want to thank Valter Cavecchia (CNR, Istituto di Fotonica e Nanotecnologie - Trento) for the suggestions and the help about openGL and for solving some troubles in my animations and applets.

Special thanks to Maria Dedò (Dipartimento di Matematica "Federigo Enriques" - Università degli Studi di Milano) for the seminary about the 4D taught to scholars in Salorno, where I saw her way of teaching that has fascinated me.

I also thank my supervisor in this work, Domenico Luminati (Laboratorio di Didattica e Comunicazione della Matematica del Dipartimento di Matematica - Università di Trento) for the enthusiasm that he has in his work and that he has transmitted to me.

Ringrazio le persone che mi hanno stimolata ad aprire gli occhi durante questi anni passati (non solo) all'università: Laura, con cui ho condiviso esperienze meravigliose, Damiana che a Tübingen è diventata una sorella, la comix che mi ha permesso di non perdere il contatto con il mondo reale anche se studiavo matematica, il mio fratellone Bruno con cui dialogo soprattutto durante le camminate in montagna, i miei genitori che non mi hanno ostacolata nelle mie scelte anche se per loro erano difficili da accettare, Giorgio che mi ha sorpresa con un regalo indimenticabile alla stazione di Bolzano.

Grazie anche agli amici conosciuti all'università, per i momenti trascorsi insieme; mi dispiace di avervi spesso solo sfiorato, senza conoscervi per bene come meritavate. In ordine sparso: Viviana, Cipro, Thomas, Rosj, Maria, Maurizio, Sara, Trentista, Giorgio, Francesca, Marianna, Silvia, Marco, Marina, Francesca, Stefania, Bolo, Demelza, Ilaria, Davide, Mattia, Robi, Bis...

Grazie ancora a chi ha fatto parte delle squadre delle Facoltiadi, in particolare quelle del 2007, per l'emozione irripetibile della doppia vittoria; al Nutria Team di dragonboat per la fatica e la prima vittoria a Borgo,

nella gara più spettacolare; alle squadrette di calcio Piallaenne e Gardolo (femminile) e alla squadra seria del Lavis, che lotta per la promozione!

Non dimentico gli amici di Tübingen, in particolare Johannes, Claudia e Alex, i Mitbewohner Yang, Irina, Daniel e Max, e gli incoscienti che hanno vissuto con me l'avventura della doppia laurea: Kirk, Giorgia, Bose, Gab, Pietro, Davide.

Bibliography

[ZS58] O. Zariski, P. Samuel,
Commutative algebra,
vol. I 1958, vol. II 1960, van Nostrand.

[Shaf77] I. R. Shafarevich,
Basic Algebraic Geometry,
Springer, 1977.

[AMd69] M. F. Atiyah, I. G. Macdonald,
Introduction to Commutative Algebra,
Addison-Wesley, 1969.

[Milne05] J. S. Milne,
Algebraic Geometry,
version 5.00, available at `http://www.jmilne.org/math/`.

[Ser56] J. P. Serre,
Géométrie algébrique et Géométrie analytique,
Annales de l'istitut Fourier, tome 6 (1956), p. 1-42.

[Hausen] J. Hausen,
Skript der Vorlesung Algebraische Geometrie,
Tübingen, Sommersemester 2006.

[Die89] J. Dieudonné,
A History of Algebraic and Differential Topology: 1900-1960,
Birkhäuser, 1989.

[Milnor78] J. W. Milnor,
Topology from the Differentiable Viewpoint,
University press of Virginia, 1978.

[BJ82] T. Bröcker, K. Jänich,
Introduction to Differential Topology,
Cambridge University Press, 1982.

[Hir76] M. W. Hirsch,
Differential Topology,
Springer, 1976.

[Kos80] R. Kosniowski,
A first Course in Algebraic Topology,
Cambridge University Press, 1980

[Spa66] E. H. Spanier,
Algebraic Topology,
Tata McGraw-Hill, 1966.

[Gre81] M. J. Greenberg, J. R. Harper,
Algebraic Topology: a First Course,
Addison-Wesley, 1981.

[Mas91] W. S. Massey,
A Basic Course in Algebraic Topology,
Springer, 1991.

[Pea75] A.R. Pears,
Dimension Theory of General Spaces,
Cambridge University Press, 1975.

[Eng77] R. Engelking,
General Topology,
Polish Scientific Publishers, 1977.

[Eng78] R. Engelking,
Dimension Theory,
North Holland Publishing Company, 1978.

[Roy62] P. Roy,
Failure of equivalence of dimension concepts for metric spaces,
Bull. Amer. Math. Soc. **68**, 1962 pag. 609-613.

[Roy68] P. Roy,
Nonequality of dimensions for metric spaces,
Trans. Amer. Math. Soc. **134**, 1968 pag. 117-132.

[Nag83] J. Nagata,
Modern Dimension Theory,
rev. and extended edition, Heldermann, 1983.

[Nag74] J. Nagata,
Modern General Topology,
2nd rev. ed. North-Holland, 1974.

[HW62] W. Hurewicz, H. Wallmann,
Dimension Theory,
rev. ed. Oxford University Press, 1962.

[Cox73] H.S.M. Coxeter,
Regular Polytopes,
Third Edition, Dover Publications, 1973.

[Ban90] T.F. Banchoff,
Beyond the Third Dimension: Geometry, Computer Graphics, and Higher Dimensions,
Freeman, 1990.

Ich erkläre hiermit, dass ich die Arbeit stelbständig und mit den angegebenen Hilfsmitteln verfaßt habe.

Trento, den 19. Dezember 2007

Ester Dalvit